

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:48:57 ; Search time 30.9427 Seconds  
(without alignments)  
1363.907 Million cell updates/sec

Title: US-09-902-563-2

Perfect score: 2378

Sequence: 1 MKLANWYLSAVLATYGYFL.....GYKSFKEAKMIRPKHFKP 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID    | Description          |
|------------|--------|---------------|--------|----------|----------------------|
| 1          | 2378   | 100.0         | 439    | 2 I37391 | fibrinogen-like pr   |
| 2          | 1853.5 | 77.9          | 432    | 2 A27447 | cytotoxic T-lympho   |
| 3          | 1849.5 | 77.8          | 432    | 2 I56934 | fibrinogen-like pr   |
| 4          | 511.5  | 21.5          | 463    | 2 A38463 | fibrinogen beta ch   |
| 5          | 506.5  | 21.3          | 312    | 2 JN0596 | fibrinogen-related   |
| 6          | 484.5  | 20.4          | 453    | 1 FGHUGB | fibrinogen gamma-B   |
| 7          | 480.5  | 20.2          | 437    | 1 FGHUG  | fibrinogen gamma-A   |
| 8          | 479.5  | 20.2          | 479    | 2 A25052 | fibrinogen beta ch   |
| 9          | 470    | 19.8          | 468    | 1 FGHUB  | fibrinogen beta ch   |
| 10         | 462    | 19.4          | 444    | 2 S05313 | fibrinogen gamma-B   |
| 11         | 461    | 19.4          | 491    | 1 FGHUB  | fibrinogen beta ch   |
| 12         | 456    | 19.2          | 438    | 2 A32670 | fibrinogen gamma c   |
| 13         | 452    | 19.0          | 334    | 2 JC5980 | fibrinogen alpha c   |
| 14         | 451    | 19.0          | 866    | 2 DA4234 | fibrinogen-related   |
| 15         | 446    | 18.8          | 282    | 2 A35084 | fibrinogen beta ch   |
| 16         | 434.5  | 18.3          | 326    | 2 B47172 | ficollin-beta - pig  |
| 17         | 431    | 18.1          | 328    | 2 A05299 | fibrinogen beta ch   |
| 18         | 428.5  | 18.0          | 323    | 2 A47172 | transforming growt   |
| 19         | 428.5  | 18.0          | 432    | 1 FGLMGS | fibrinogen gamma c   |
| 20         | 423    | 17.8          | 326    | 2 S61517 | ficollin-l precursor |
| 21         | 423    | 17.8          | 1353   | 1 JH0675 | restrictin precurs   |
| 22         | 417.5  | 17.6          | 1356   | 2 A45445 | janusin precursor,   |
| 23         | 404    | 17.0          | 437    | 1 FGHUGA | fibrinogen gamma-A   |
| 24         | 404    | 17.0          | 445    | 1 FGHUGB | fibrinogen gamma-B   |
| 25         | 403.5  | 17.0          | 220    | 2 S28170 | fibrinogen homolog   |
| 26         | 403.5  | 17.0          | 417    | 2 S65944 | tenascin-X - pig (   |
| 27         | 401.5  | 16.9          | 4135   | 2 T42629 | tenascin-X - bovin   |
| 28         | 397.5  | 16.7          | 3566   | 1 A40701 | tenascin-X precurs   |
| 29         | 396    | 16.7          | 2019   | 1 JQ1322 | tenascin precursor   |

30 393.5 16.5 860 2 I48839  
31 393.5 16.5 4006 2 T09070  
32 388 16.3 2201 2 A32160  
33 386.5 16.3 1746 1 S19694  
34 386 16.2 1810 1 A32230  
35 382.5 16.1 1914 2 T42635  
36 356.5 15.0 641 1 A41932  
37 346.5 14.6 774 2 A39832  
38 215.5 9.1 463 2 T15876  
39 198 8.3 915 2 T21773  
40 198 8.3 927 2 T21772  
41 184.5 7.8 127 2 PC2036  
42 160 6.7 146 2 T32255  
43 160 6.7 431 2 T29850  
44 160 6.7 452 2 T26827  
45 140 5.9 933 2 A31930

## ALIGNMENTS

### RESULT 1

I37391

fibrinogen-like protein expressed in T lymphocytes (PT49) - human

C:Species: Homo sapiens (man)

C>Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 21-Jul-2000

C:Accession: I37391; S47273

R:Ruegg, C.; Pytel, R.

Gene 160, 257-262, 1995

A:Title: Sequence of a human transcript expressed in T-lymphocytes and encoding a fib

A:Reference number: I37391; MUID:95369700; PMID:7642106

A:Accession: I37391

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-439 <RES>

A:Cross-references: EMBL:Z36531; NID:9535184; PIDN:CA85298.1; PID:9535185

A:Note: submitted to the EMBL Data Library, August 1994

C:Superfamily: fibrinogen gamma chain; fibrinogen gamma/gamma homology

F:210-435/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 100.0%; Score 2378; DB 2; Length 439;  
Best Local Similarity 100.0%; Pred. No. 4.3e-157;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLANWYLSAVLATYGYFLVYANNTEELKDRADKCPVRLSRGKCEAGECPYVS 60  
DB 1 MKLANWYLSAVLATYGYFLVYANNTEELKDRADKCPVRLSRGKCEAGECPYVS 60  
QY 61 LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQCKLQADDNGDGRNGLLPSTG 120  
DB 61 LPPLTIQLPKQFSRIEEVFKEVQNLKEIVNSLKKSCQCKLQADDNGDGRNGLLPSTG 120  
QY 121 APGEVGNRRVRELESVKNLSSELNKAKEINVHGRLEKLNLYNMNNIENVYDSKVANL 180  
DB 121 APGEVGNRRVRELESVKNLSSELNKAKEINVHGRLEKLNLYNMNNIENVYDSKVANL 180  
QY 181 TFVYNSLDGKSCPSQEQIQSRPVQHLYIKDCSDYIAIGKRSSETYRVTPDPKNSFEV 240  
DB 181 TFVYNSLDGKSCPSQEQIQSRPVQHLYIKDCSDYIAIGKRSSETYRVTPDPKNSFEV 240  
QY 241 YCDMETGGGWTVLQARLDGSTNFTRTWQYKAGFNLRRERFNLGNOKIHLLTKSEKML 300  
DB 241 YCDMETGGGWTVLQARLDGSTNFTRTWQYKAGFNLRRERFNLGNOKIHLLTKSEKML 300  
QY 301 RIDLEDFENGVELYALYDQFYVANEFLKRLHVGNNGTAGDALRFNKHYNHLKFFTPD 360  
DB 301 RIDLEDFENGVELYALYDQFYVANEFLKRLHVGNNGTAGDALRFNKHYNHLKFFTPD 360  
QY 361 KNDRYFSGNCGLYYSSGWFDFACLSANLNGKYHOKYEGVNGTGWGTPGVSEAHFGG 420  
DB 361 KNDRYFSGNCGLYYSSGWFDFACLSANLNGKYHOKYEGVNGTGWGTPGVSEAHFGG 420  
QY 421 YKSFKEAKMIRPKHFKP 439

tenascin-X - mouse  
probable tenascin  
tenascin-C - human  
tenascin precursor  
tenascin precursor  
tenascin Y precurs  
fibrinogen alpha-I  
scabrous locus (sc  
hypothetical prote  
hypothetical prote  
hypothetical prote  
microfibril-assoc  
hypothetical prote  
hypothetical prote  
cytotactin - chick



```
QY 125 SEVNLKSLKNAKEEINVLHGRLEKLNVLNNNNTENYVSKVANLTFVNSLDGKCSK 194
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 160 NLDNNIPSSLRVLRVAVLDSLHKKIQKL-----ENATATQT-----DYCRSPC 201

QY 195 PSQEQIQSPVOHLLYKDCSDYAIKRSSEYRVTPDKNSSEFVYCDMETGCGWTVL 254
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 202 -----VASCNIPVSGREEDIRKGGTSEMYIIQDPFTTPYRYVCDMETDNGCWTLI 256
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 255 QARLDGSTFTRTWDYKAGFGNLR-----EFWLGNDRKTHLLTKSKEMILRID 303
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 257 QNRQDSVNFGRVDFVGRVGRVGRVGRVGRVGRVGRVGRVGRVGRVGRVGRVGRV 316
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 304 LEQFNGVELYALYDOFYVANEFLKYLRLHVGNYNGTAGDALR--FNKHYH-----HDLK 354
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 317 MEWNGDKVSALYGGFTTHNEGKYLQVSNYKGNAGNALMEGASOLQGENRTMTHNGM 376
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 355 FTTPTDKNDRY----PSGNCGLYSSGWDFDACLSANLNGKYY-----HQYRGVNRN 403
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 377 YFTYDRDNDGLTTPDKQCKEDGGGWYNRCHAANPNRYGGTYSDNMAKHGTDD 436
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

404 GTFWGTGPGVSEAHFPGYKSPKAKMMLRP 434
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
437 GIVVMNKG-----SWY--SMKMSMKIKP 459

RESULT 5
JN0596
fibrinogen-related protein HFREP-1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: JN0596
R:Yamamoto, T.; Gotoh, M.; Sasaki, H.; Terada, M.; Kitajima, M.; Hirohashi, S.
Biochem. Biophys. Res. Commun. 193, 681-687, 1993
A:Title: Molecular cloning and initial characterization of a novel fibrinogen-related gene
A:Reference number: JN0596; MUID:93290661; PMID:8390249
A:Accession: JN0596
A:Molecule type: mRNA
A:Residues: 1-312 <YAM>
A:Cross-references: GB:D14446; NID:q393314; PIDN:BA0A3336.1; PID:g9393315
A:Experimental source: liver
C:Superfamily: fibrinogen beta/gamma homology
F:1-17/Domain: signal sequence status predicted <SIG>
F:18-312/Product: fibrinogen-related protein HFREP-1 #status predicted <MAT>
F:80-305/Domain: fibrinogen beta/gamma homology <PBG>

Query Match 21.3%; Score 506.5; DB 2; Length 312;
Best Local Similarity 36.8%; Pred. No. 1e-27;
Matches 118; Conservative 38; Mismatches 108; Indels 57; Gaps 9;

129 RVRELESEVNLKSLKNAKEEINVLHGRLEKLNVLNNNNTENYVSKVANLTFVNSLD 188
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 36 QVRLLTRVKKQVKKIKQLQE-----NEVQFLDKGDETVVD----- 73

QY 189 GKSKCPSEQIQSPVOHLLYKDCSDYAIKRSSEYRVTPDKNSSEFVYCDMETWG 248
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 74 -----LGSKRQ-----YADCEIFNDGYKLSGFKIRPLOSFAEFVSDMSD-G 117

QY 249 GGTVLQARLDGSTFTRTWDYKAGFGNLR-----RREFWLGNDKTHLLTKSKEMILRIDLE 305
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 118 GGTVLQARLDGSTFTRTWDYKAGFGNLR-----RREFWLGNDKTHLLTKSKEMILRIDLE 305
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 306 DFNGVELYALYDOFYVANEFLKYLRLHVGNYNGTAGDALRFNKH-----YHDLKFTTTP 359
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 178 DFEKNSRVAQYKNFVGVDEKFNELNTEYSGTAGDSLACNFPHEVQWASHQRMKFTW 237

QY 360 DKNDRYPSGNCGLYSSGWDFDACLSANLNGKYYHQYRG--VRNGIPWGTWPCVSEAHF 418
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 238 DRDHNY--EGNCAEEDSGWMTFNCHSALNNGVYSGPYTAKTDNGIWWYTW----- 288

QY 419 GGYKSKFEAKMMLRPKHFP 439
:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 289 HGHWYSLKSVVWKLRPNDP 309
```

## RESULT 6

FEHUGB

fibrinogen gamma-B chain precursor [validated] - human  
N:Alternate names: coagulation factor I; fibrinogen gamma-55 chain  
C:Species: Homo sapiens (man)  
C:Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 08-Dec-2000  
C:Accession: A90494; A92448; A90453; A28203; B28203; I37390; A03126  
R:Rixon, M.W.; Chung, D.W.; Davies, E.W.  
Biochemistry 24, 2077-2086, 1985  
A:Title: Nucleotide sequence of the gene for the gamma chain of human fibrinogen.  
A:Reference number: A90494; MUID:85252774; PMID:2990550  
A:Accession: A90494  
A:Molecule type: DNA  
A:Residues: 1-113, 'I', 115-453 <RIX>  
A:Cross-references: GB:M10014; GB:J00134; GB:J00135; GB:X00086; NID:q182438; PIDN:RAB  
J:Forname Jr., A.J.; Cummings, D.E.; Comeau, C.M.; Kant, J.A.; Crabtree, G.R.  
J. Biol. Chem. 259, 12826-12830, 1984  
A:Title: Structure of the human gamma-fibrinogen gene. Alternate mRNA splicing near t  
A:Reference number: A92448; MUID:85030379; PMID:6092346  
A:Accession: A92448  
A:Molecule type: DNA  
A:Residues: 286-453 <FOR>  
R:Wolfenstein-Todel, C.; Mosesson, M.W.  
Biochemistry 20, 6146-6149, 1981  
A:Title: Carboxy-terminal amino acid sequence of a human fibrinogen gamma-chain varia  
A:Reference number: A90453; MUID:82068993; PMID:7306501  
A:Accession: A90453  
A:Molecule type: protein  
A:Residues: 411-434, 'Y', 436-440, 'Z', 442, 'Z', 444, 'B', 446-447, 'R', 449, 'ZBB', 453 <WOL>  
R:Francis, C.W.; Mueller, E.; Henschen, A.; Simpson, P.J.; Marder, V.J.  
Proc. Natl. Acad. Sci. U.S.A. 85, 3358-3362, 1988  
A:Title: Carboxyl-terminal amino acid sequences of two variant forms of the gamma-chain  
A:Reference number: A94194; MUID:88217900; PMID:3368448  
A:Accession: A94194  
A:Molecule type: protein  
A:Residues: 433-449 <FRA>  
A:Accession: B28203  
A:Molecule type: protein  
A:Residues: 433-453 <FR2>  
R:Marchetti, L.; Zanelli, T.; Malcovati, M.; Tenchini, M.L.  
DNA Seq. 1, 419-422, 1991  
A:Title: Polymorphism of the human gamma chain fibrinogen gene.  
A:Reference number: I37390; MUID:92119334; PMID:1685103  
A:Accession: I37390  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 75-286 <RES>  
A:Cross-references: EMBL:X51473; NID:q31410; PIDN:CAA35837.1; PID:g930064  
C:Comment: The two forms of gamma chain, A (see PIR:FGHUG) and B, arise by alternate  
ntron, which makes this chain different from the gamma-B chain at positions 434-437 a  
C:Comment: The gamma-B chain is present in about 10% of the fibrinogen molecules in p  
C:Genetics:  
A:Gene: GDB:FGG  
A:Cross-references: GDB:119132; OMIM:134850  
A:Map position: 4q28-4q28  
A:Introns: 26/3; 41/3; 103/1; 134/2; 178/1; 222/3; 284/2; 377/1  
C:Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR  
ins are contained in the core. Two three-chain coiled coils emerge from this core and  
from the distal domain nodes.  
C:Function:  
A:Description: fibrinogen cleaved by thrombin yields monomers that are polymerized in  
A:Pathway: blood coagulation  
C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology  
C:Keywords: alternative splicing; blood coagulation; calcium; coiled coil; glycoprote  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-453/Product: fibrinogen gamma-B chain #status experimental <MPT>  
F:176-415/Domain: fibrinogen beta/gamma homology <PBG>  
F:341-355/Domain: calcium binding #status predicted <CAB>  
F:400-422/Region: polymerization site, binding to the amino end of the alpha chain of  
F:34/disulfide bonds: interchain (to gamma-35) #status predicted  
F:35/disulfide bonds: interchain (to gamma-34) #status predicted  
F:45/disulfide bonds: interchain (to beta-110) #status predicted



|  |     |                  |  |     |
|--|-----|------------------|--|-----|
| Db   | 158 | EAQCQE-PCKDTVQ   | ---IHDITGCKODDIANKGAKOSGLYFTKPLKANQOFLVYCEIDGS         | 213 |
| QY   | 248 | GGGWTVLQARLDG    | STNFTRWODYKAGECNL-----RREFWLGNDKIHLTKSKEM--ILR         | 301 |
| Db   | 214 | NGWTVFOKRLDGS    | VDFEKNWIKYQEGFCHLSPGTTTFWPLGNEKIHLLISTQSAIPALR         | 273 |
| QY   | 302 | IDLEDENGVELYALD  | YDYFYANFELKYRLHVGNY-NCTAGDAL-RFNKHYNHIDLKFT--          | 357 |
| Db   | 274 | VELEDWNGTSTADY   | AMFKVGPBARKYRLTYAYFAGDGADGADFDFGDDPSDKFTFSH            | 333 |
| QY   | 358 | -----TPDKNDRYP   | SGNCLYISGWMFADCLASLANLKYTH-----QKRYGRVNR               | 404 |
| Db   | 334 | NGMOfSTWdNDKf    | -EGNCAeDgSGWMNkCHAGHLNcVYOGgTYSKASTPNCgYDNG            | 392 |
| QY   | 405 | IFWCTW           | 410  |     |
| Db   | 393 | IITATW           | 398  |     |
| RESULT 8   |     |                  |  |     |
| A25052   |     |                  |  |     |
| fibrinogen beta chain - sea lamprey (fragments)                              |     |                  |  |     |
| N-Contans: fibrinopeptide B  |     |                  |  |     |
| C-Species: Petromyzon marinus (sea lamprey)                                  |     |                  |  |     |
| C.Date: 25-Oct-1987 sequence_revision 19-Feb-1999 #text_change 13-Aug-1999   |     |                  |  |     |
| A-Accession: A25052; A03124; B03124  |     |                  |  |     |
| R:Bohoun, V.L.; Doolittle, R.F.; Pontes, M.; Strong, D.D.                    |     |                  |  |     |
| Biochemistry 25, 6512-6516, 1986   |     |                  |  |     |
| A>Title: Complementary DNA sequence of lamprey fibrinogen beta chain.        |     |                  |  |     |
| A:Reference number: A25052; MUID:87075682; PMID:3790537                      |     |                  |  |     |
| A:Accession: A25052  |     |                  |  |     |
| A:Molecule type: mRNA  |     |                  |  |     |
| A:Residues: 39-479 <BO>  |     |                  |  |     |
| A:Cross-references: GB:MI4773; NTD:g213191; PIDN:AA49261.1; PID:g213192      |     |                  |  |     |
| R:Contrell, B.A.; Doolittle, R.F.  |     |                  |  |     |
| Biochim. Biophys. Acta 453, 426-438, 1976                                    |     |                  |  |     |
| A>Title: Amino acid sequences of lamprey fibrinopeptides A and B and charac  |     |                  |  |     |
| A:Reference number: A03120; MUID:77065679; PMID:999898                       |     |                  |  |     |
| A-Accession: A03124  |     |                  |  |     |
| A:Molecule type: protein   |     |                  |  |     |
| A:Residues: 1-36 <COT>   |     |                  |  |     |
| A-Accession: B03124  |     |                  |  |     |
| A:Molecule type: protein   |     |                  |  |     |
| A:Residues: 37-42 <COT2>   |     |                  |  |     |
| C-Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrin |     |                  |  |     |
| C:Keywords: blood coagulation; glycoprotein; sulfoliprotein                  |     |                  |  |     |
| F:1-36/Product: fibrinopeptide B #status experimental <FPB>                  |     |                  |  |     |
| F:37-479/Product: fibrin beta chain #status experimental <MAT>               |     |                  |  |     |
| F:90-219/Domain: fibrinogen disulfide ring homology <FDR>                    |     |                  |  |     |
| F:229-477/Domain: fibrinogen beta/gamma homology <FPG>                       |     |                  |  |     |
| F:7/Binding site: sulfate (Tyr) (covalent) #status experimental              |     |                  |  |     |
| F:7/Binding site: carbohydrate (Asn) (covalent) #status experimental         |     |                  |  |     |
| Query Match 20.2%; Score 479.5; DB 2; Length 479;                            |     |                  |  |     |
| Best Local Similarity 29.2%; Pred. No. 1.3e-25;                              |     |                  |  |     |
| Matches 131; Conservative 66; Mismatches 138; Indels 113; Gaps               |     |                  |  |     |
| QY   | 29  | EIKDERAKDVCPRVLE | SRGKCEAGECPQVLSPLTIQLPKQFSRIEEVFKVQNLKEI               | 88  |
| Db   | 102 | ELREELLKQDRV     | -----YKISMLK-----ONLTYF                                | 128 |
| QY   | 89  | VNSLKSKSCDCKLQ   | ADDDGPGRNGLLLFSTGAPCEVGNRVRELESEVNKLSELKNK             | 148 |
| Db   | 129 | INSFDR           | -----MASDSNTLKQN-----VOTLRRLNLSRSGSTHVAQ               | 164 |
| QY   | 149 | BEINVLHGRLEKLN   | VNNNIENYVDKSVANLT FVNSLDGKCSK----CPSEQEQTSPR           | 204 |
| Db   | 165 | KEI              | ---ENRYKEVKI---RIESTVAGSLRSMKSVLHEDLRAMQWEAIKTOKEJCSAP | 217 |
| QY   | 205 | -----VOHLIYKDC   | SDYYAIGKRSSERYTTPDPKNKSFEYVCDMTGGMGTVLQARL             | 258 |
| Db   | 218 | CTVNCRVPVVS      | SGMHCEDIRNGGRTSEAYIOPDLFEPYKFCVDMESHGGMVTVONRV         | 277 |

Best Local Similarity 30.6%; Pred. No. 5.9e-25;  
Matches 129; Conservative 57; Mismatches 149; Indels 86; Gaps 13;

QY 61 LPPLTIQLPKQFSRIE-VFKEVONLKEIVNSLKSKQCDC--KLQADONGDPGRNGLLLP 117  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 82 LCPTCKLQDTLVRQRPPIKRSIEDLRNTVDVSRTSSSTFFQVITLLKNMKGR----- 135

QY 118 STGAPGEVDNRVRELESEVNKLSELNAKEEINVLHGFELEKLNLVNWNNIENYVDSKV 177  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 136 -----NQVODNNENVNYESS-----HLEKHLYIDETVKNNIPTKL 172

QY 178 ANLTFWNSLDGKSCKPSEQEIQR-----PVQHLYIKDCSDYYAIGKRSE 225  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 173 RVLRSILENLRSKIQLKESDVSTQMEYCRTPCTVTNCIPV--VSGCEKEKLIIRNEGTSSE 230

QY 226 TVRVTPDPKNSFEVCMDNETMGGGWTVLAARDGDSNTRTWQDYKAGFCGNLRR----- 280  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 231 WYLQPEDSKPRVYCDMKTEKGWVIQNROGSUDFGRKWDYPKOGFNATNAEGK 290

QY 281 -----EFWLGNDKTHLLTKSKEMILRIDLEDPNGVELYLDQFYVANEFKYRLHW 333  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 291 KYCGVPGEYWLGNDRISOLTNGPPTKLLIEMDWKGDVKTALYEGFTVQNEANKYQLSVS 350

QY 334 NVNQTAGDAL-----RENKHYN-HDLKFFTTPKDNDRY----PSGNCGLYSYGKW 380  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 351 KYKGTAGNALLEGASOLVGENTWTIHNSFFEFTYDRDNDGHKTTDPRKQSKEDGGGNW 410

QY 381 FPAULSANLNGKY-----HQKYRVRNGIFWGTPGVSEAHPGGYKSFKAKMWIR 433  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 411 YNRCHAANPNRGYYWGAYTWMAMKHGTDGGVVMNWQG-----SWY--SMKKMSMKIR 462

QY 434 P 434  
|  
Db 463 P 463

RESULT 10  
S05313  
fibrinogen gamma-B chain precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 13-Aug-1999  
C:Accession: S05313  
R:Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.  
Nucleic Acids Res. 17, 6397, 1989  
A:Title: Nucleotide and deduced amino acid sequence of a gamma subunit of bovine fibrinogen  
A:Reference number: S05313; MUID:89366676; PMID:2771651  
A:Accession: S05313  
A:Molecule type: mRNA  
A:Residues: 1-444 <BROW>  
A:Cross-references: EXBL:X15556; NID:g349; PIDN:CAA33562.1; PID:g350  
A>Note: the authors translated the codon ACT for residue 105 as Ala and ATT for residue 106 as Asn  
C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-444/Product: fibrinogen gamma-B chain #status predicted <MAT>  
F:174-414/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 19.4%; Score 462; DB 2; Length 444;  
Best Local Similarity 30.5%; Pred. No. 2e-24;  
Matches 142; Conservative 51; Mismatches 164; Indels 108; Gaps 19;

QY 9 LSSAVLATYGLVANNETEIEIKDERAKDVCVRLESRGKCEEAGECPYOVSLPPTIQL 68  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 18 LSSACLA-----YVATRNCCLIDEREGSYCPT-----TCGIADFNNYTQSV 60

QY 69 PKQFSRIEEVFKEVONLKEIVNSLKSKQCDCQLQADDNGDPCRNGLLLPSTGAPGEVDGN 128  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 61 DKDLRTLLEGILYOVENTKTSSEBELVKA-----IQTSYNPDQ----- 96

QY 129 RVRELESEVNKLSELNAK---BEINVHLGRLEKLNLVNWNNIENYVDSKV 177  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 97 -----PSKPNNTESATKNSKMMEIE-----WKYETLISTHETSFLEQEVYNSQSI 145

QY 178 ANLTFWNSLDGKSCKPSEQEIQRVQHLIVKDCSDYYAIGKRSETRYVTPDPKNS 237

Db 146 VNLKRVVLEANCQE-PCQDIVK---IHDVTGDCQDVANKAKESCLYFIRP-LKAKQ 200  
QY 238 FEWYCDMTGCGKWTVLQRLDGSNTFTWQDYKAGFNL-----RRFWLGNDKIHL 291  
Db 201 FUYCEIDSGNGWTFQRLDGSUDFKKNWQYKCFGHLSPTGTGNTFWLGNKEIHL 260  
QY 292 LTKSEM---ILRLDLEDFNGVELYALYDQFVANEFLKYLRLHVGNY-NGTAGDAL----- 343  
Db 261 ISTQSSIPYVLRIOLEDWNGRSTADYASFVKVTCENDKYRLFYAYFTGGDAGDAFGDYDF 320  
QY 344 ---RNNKYH-NHDLKFTTPDKNDRPSNGGLYSSWQWDFDACLGSANLNGKYH----- 395  
Db 321 GDDSDKFTFTHNGMOFSTWSDNDKY-DGNCAEQVIGICWNNKCHAGLNLGVYQGGTY 379  
QY 396 ---QKRVGRNGIFWGTGCVSEAHPPGGYKS---SFKKAKMMIRP 434  
Db 380 SKTSTPNGYDNGLIWATW-----KSRWYSMKKTKTIIP 413

SULT 11  
HUB  
Fibrinogen beta chain precursor [validated] - human  
N:Alternate names: coagulation factor I  
N:Contains: fibrinopeptide B  
C:Species: Homo sapiens (man)  
C:Date: 24-Apr-1984 #sequence\_revision 31-Mar-1993 #text\_change 08-Dec-2000  
C:Accession: B43568; A90469; B90469; I37389; A94433; A90437; A94309; G54223; A03121; B37  
R:Chung, D.W.; Harris, J.E.; Davies, E.W.  
Adv. Exp. Med. Biol. 281, 39-48, 1990  
A:Title: Nucleotide sequences of the three genes coding for human fibrinogen.  
A:Reference number: A43568; UID:91344740; PMID:2102623  
A:Accession: B43568  
A:Molecule type: DNA  
A:Residues: 9-191, 'P', 193-491 <CHD>  
R:Chung, D.W.; Que, B.G.; Rixon, M.W.; Mace Jr., M.; Davie, E.W.  
Biochemistry 22, 3244-3250, 1983  
A:Title: Characterization of complementary deoxyribonucleic acid and genomic deoxyribonu  
A:Reference number: A90469; UID:83283433; PMID:6688356  
A:Accession: A90469  
A:Molecule type: DNA  
A:Residues: 1-38 <CHL>  
A:Accession: B90469  
A:Molecule type: mRNA  
A:Residues: 9-191, 'A', 193-491 <CH2>  
A:CROSS-references: GB:J00129; MID:g182429; PIDN:AA52429.1; PID:g182430  
R:Huber, P.; Dalmon, J.; Courtois, G.; Laurent, M.; Assouline, Z.; Marguerie, G.  
Nucleic Acids Res. 15, 1615-1625, 1987  
A:Title: Characterization of the 5'-flanking region for the human fibrinogen beta gene.  
Reference number: I37389; UID:87146483; PMID:3029722  
Accession: I37389  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-38 <HUB>  
A:CROSS-references: EMBL:X05018; NID:g31400; PIDN:CAA28674.1; PID:g31401  
R:Henschen, A.; Lottspeich, F.; Southan, C.; Tøpfer-Petersen, E.  
In: Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Pe  
A:Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural v  
A:Reference number: A94433  
A:Contents: carbohydrate binding  
A:Accession: A94433  
A:Molecule type: protein  
A:Residues: 31-137, 'QS', 140-144, 'QF', 147-491 <HEN>  
R:Watt, K.W.K.; Takagi, T.; Doolittle, R.F.  
Biochemistry 18, 68-76, 1979  
A:Title: Amino acid sequence of the beta chain of human fibrinogen.  
A:Reference number: A90437; UID:79124640; PMID:420779  
A:Accession: A90437  
A:Molecule type: protein  
A:Residues: 31-144, 'QF', 147-231, 'D', 233-330, 'E', 332-491 <WAT>  
R:Blomback, B.; Hessel, B.; Hogg, D.  
Thromb. Res. 8, 639-658, 1976  
A:Title: Disulfide bridges in NH-2-terminal part of human fibrinogen.

A:Reference number: A94309; UID:76225080; PMID:936108  
A:Contents: disulfide bonds  
A:Accession: A94309  
A:Molecule type: protein  
A:Residues: 31-112, 'E', 114-137, 'QS', 140-144, 'QF', 147-148 <BLO>  
R:Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.  
Biochemistry 33, 1988-1993, 1994  
A:Title: Identification of proteins associated with apolipoprotein A-I-containing lip  
A:Reference number: A54223; UID:94162201; PMID:8117655  
A:Accession: G54223  
A:Molecule type: protein  
A:Residues: 164-174 <KUN>  
A:Note: Identification of tryptic peptides from high-density lipoproteins  
R:Henschen, A.; Lottspeich, F.; Kehl, M.; Southan, C.  
Ann. N. Y. Acad. Sci. 408, 28-43, 1983  
A:Title: Covalent structure of fibrinogen.  
A:Reference number: A90037; UID:83254370; PMID:6575689  
R:Gardlund, B.; Hessel, B.; Marquerie, G.; Murano, G.; Blomback, B.  
Eur. J. Biochem. 77, 595-610, 1977  
A:Title: Primary structure of human fibrinogen. Characterization of disulfide-contains  
A:Reference number: A91249; UID:77245999; PMID:891553  
A:Contents: annotation; disulfide bonds  
R:Doolittle, R.F.; Takagi, T.; Watt, K.; Bouma III, H.; Cottrell, B.A.; Cassman, K.G.  
In: Regulatory Proteolytic Enzymes and Their Inhibitors, Magnusson, S., Ottesen, M., F  
A:Title: The structures of fibrinogen and fibrin.  
A:Reference number: A94437  
A:Contents: annotation; disulfide bonds  
R:Doolittle, R.F.  
Annu. Rev. Biochem. 53, 195-229, 1984  
A:Title: Fibrinogen and fibrin.  
A:Reference number: A90041; UID:84305751; PMID:6383194  
A:Contents: annotation; review, EM structure, polymerization, ligands  
R:Chung, D.W.; Rixon, M.W.; Que, B.G.; Davie, E.W.  
Ann. N. Y. Acad. Sci. 408, 449-456, 1983  
A:Title: Cloning of fibrinogen genes and their cDNA.  
A:Reference number: A90038; UID:83254384; PMID:6575700  
A:Contents: annotation  
R:Kirschbaum, N.E.; Budzynski, A.Z.  
J. Biol. Chem. 265, 13669-13676, 1990  
A:Title: A unique proteolytic fragment of human fibrinogen containing the Aalpha COOH  
A:Reference number: A37117; UID:90337977; PMID:2143188  
A:Contents: annotation; hemostatic cleavage site  
A:Note: Hementin, a protease from Haemotelia ghilianii, the giant South American lee  
C:Comment: The conversion of fibrinogen to fibrin is triggered by thrombin, which cle  
ization sites responsible for the formation of the soft clot.  
C:Comment: The soft clot is converted into the hard clot by factor XIIIa (fibrin-stab  
ger) and between alpha chains (weaker) of different monomers.  
C:Comment: All fibrinogen chains are synthesized in the liver.  
C:Genetics:  
A:Gene: GDB:FCB  
A:CROSS-references: GDB:119130; OMIM:134830  
A:Map position: 4q28-4q28  
A:Introns: 38/3; 102/3; 164/1; 240/1; 278/1; 320/1; 415/2  
C:Complex: The fibrinogen molecule is a hexamer containing two sets of alpha (see PIR  
ins are contained in the core. Two three-chain coiled coils emerge from this core and  
from the distal domain nodes.  
C:Function:  
A:Description: fibrinogen cleaved by thrombin yields monomers that are polymerized in  
A:Pathway: blood coagulation  
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disu  
C:Keywords: blood coagulation; coiled coil; glycoprotein; liver; plasma; pyroglyutamic  
F:1-30/Domain: (or 4-30 or 15-30) signal sequence #status experimental <SIG>  
F:31-491/Product: fibrinogen beta chain #status experimental <WAT>  
F:31-44/Product: fibrinopeptide B #status experimental <APT>  
F:45-491/Product: fibrin beta chain #status experimental <FCB>  
F:45-47/Region: polymerization site  
F:99-228/Domain: fibrinogen disulfide ring homology <FDR>  
F:238-487/Domain: fibrinogen beta/gamma homology <FBG>  
F:31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi  
F:44-45/Cleavage site: Arg-Gly (thrombin) #status experimental  
F:95/Disulfide bonds: interchain (to alpha-55) #status experimental  
F:106/Disulfide bonds: interchain (to alpha-68) #status experimental





Mon Dec 2 07:10:23 2002

us-09-902-563-2.rpr

Page 10

Db 241 LLSNLNGQYY--DYSGAPS-IYWSYLPDNDQIP-----FAEMKLRNR 280

Search completed: December 2, 2002, 06:54:01  
Job time : 32.9427 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:41:25 : Search time 36,7445 Seconds  
(without alignments)  
495,533 Million cell updates/sec

Title: US-09-902-563-2

Perfect score: 2378

Sequence: 1 MKLANWYWLSSAVLATYGLF.....GYKSSEKAKMIRPKFKP 439

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID        | Description         |
|------------|--------|---------------|--------|--------------|---------------------|
| 1          | 2378   | 100.0         | 439    | 1 FGL2_HUMAN | Q14314 homo sapien  |
| 2          | 1853.5 | 77.9          | 432    | 1 FGL2_MOUSE | P12804 mus musculus |
| 3          | 525    | 22.1          | 496    | 1 AGP2_BOVIN | O35088 mus musculus |
| 4          | 524.5  | 22.1          | 375    | 1 AGP2_BOVIN | O77802 bos taurus   |
| 5          | 522    | 22.0          | 496    | 1 AGP2_HUMAN | O15123 homo sapien  |
| 6          | 511.5  | 21.5          | 463    | 1 FIBB_CHICK | O02020 gallus gall  |
| 7          | 506.5  | 21.3          | 312    | 1 FGL1_HUMAN | O08830 homo sapien  |
| 8          | 486.5  | 20.5          | 498    | 1 AGP1_HUMAN | O15389 homo sapien  |
| 9          | 485.5  | 20.4          | 493    | 1 ANL2_MOUSE | O9r045 mus musculus |
| 10         | 484.5  | 20.4          | 453    | 1 FIBG_HUMAN | P02679 homo sapien  |
| 11         | 483.5  | 20.3          | 498    | 1 AGP1_MOUSE | O08538 mus musculus |
| 12         | 480.5  | 20.2          | 493    | 1 ANL2_HUMAN | Q9uk09 homo sapien  |
| 13         | 479.5  | 20.2          | 477    | 1 FIBB_PETMA | P02678 petromyzon   |
| 14         | 471    | 19.8          | 468    | 1 FIBB_BOVIN | P02676 bos taurus   |
| 15         | 470.5  | 19.8          | 509    | 1 AGP4_MOUSE | O9wh66 mus musculus |
| 16         | 467    | 19.6          | 481    | 1 AGP1_BOVIN | O18920 bos taurus   |
| 17         | 462    | 19.4          | 444    | 1 FIBG_BOVIN | P12799 bos taurus   |
| 18         | 461    | 19.4          | 451    | 1 FIBB_HUMAN | P02675 homo sapien  |
| 19         | 456    | 19.2          | 438    | 1 FIBG_XENLA | P17634 xenopus lae  |
| 20         | 454.5  | 19.1          | 479    | 1 FIBB_RAT   | P14480 rattus norv  |
| 21         | 452    | 19.0          | 334    | 1 FCN1_MOUSE | O70165 mus musculus |
| 22         | 451    | 19.0          | 866    | 1 FIBA_HUMAN | P02671 homo sapien  |
| 23         | 450    | 18.9          | 503    | 1 AGP4_HUMAN | O9y264 homo sapien  |
| 24         | 446    | 18.8          | 282    | 1 FIBA_PARPA | P19477 parastichop  |
| 25         | 441    | 18.5          | 782    | 1 FIBA_RAT   | P06399 rattus norv  |
| 26         | 438.5  | 18.4          | 335    | 1 FCN1_RAT   | O9wts8 rattus norv  |
| 27         | 428.5  | 18.0          | 432    | 1 FIBG_PETMA | P04115 petromyzon   |
| 28         | 428    | 18.0          | 319    | 1 FCN2_RAT   | P57756 rattus norv  |
| 29         | 427.5  | 18.0          | 741    | 1 FIBA_CHICK | P14448 gallus gall  |
| 30         | 427    | 18.0          | 326    | 1 FCN1_HUMAN | O00602 homo sapien  |
| 31         | 421.5  | 17.7          | 306    | 1 FCN2_MOUSE | O70497 mus musculus |
| 32         | 420    | 17.7          | 313    | 1 FCN2_HUMAN | Q15485 homo sapien  |
| 33         | 419    | 17.6          | 445    | 1 FIBG_RAT   | P02680 rattus norv  |

#### RESULT 1

| ID | FGL2_HUMAN  | STANDARD | PRT | 439 AA |
|----|---|----------|-----|--------|
| AC | Q14314  |          |     |        |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update)   |          |     |        |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update)   |          |     |        |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update)   |          |     |        |
| DE | Fibrinogen precursor (Fibrinogen-like protein 2) (PT49).  |          |     |        |
| GN | FGL2.   |          |     |        |
| OS | Homo sapiens (Human).   |          |     |        |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |          |     |        |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.   |          |     |        |
| OX | NCBI_TaxID=9606;  |          |     |        |
| RN | [1]   |          |     |        |
| RP | SEQUENCE FROM N.A.  |          |     |        |
| RC | TISSUE=Small intestine;   |          |     |        |
| RX | MEDLINE=95369700; PubMed=7642106;   |          |     |        |
| RA | Ruegg C., Pytel R.;   |          |     |        |
| RT | "Sequence of a human transcript expressed in T-lymphocytes and encoding a fibrinogen-like protein.";  |          |     |        |
| RL | Gene 160:257-262(1995).   |          |     |        |
| RN | [2]   |          |     |        |
| RP | SEQUENCE FROM N.A.  |          |     |        |
| RA | Yuvaraj S., Liu M., Marsden P., Levy G.;  |          |     |        |
| RT | "Cloning and characterization of Hfg12: the human counterpart to the mouse gene Fgl2.";   |          |     |        |
| RL | Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.   |          |     |        |
| RN | [3]   |          |     |        |
| RP | SEQUENCE FROM N.A., AND VARIANT GL0-53.   |          |     |        |
| RA | Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Ozuna M., Yi Q., Nickerson D.A.;  |          |     |        |
| RL | Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.   |          |     |        |
| RN | [4]   |          |     |        |
| RP | CHARACTERIZATION.   |          |     |        |
| RX | MEDLINE=98309432; PubMed=9647217;   |          |     |        |
| RA | Marazzi S., Blum S., Hartmann R., Gundersen D., Schreyer M.,  |          |     |        |
| RT | "Characterization of human fibrinogen, a fibrinogen-like protein secreted by T lymphocytes.";   |          |     |        |
| RL | J. Immunol. 161:138-147(1998).  |          |     |        |
| CC | -1- FUNCTION: MAY PLAY A ROLE IN PHYSIOLOGIC LYMPHOCYTE FUNCTIONS AT MUCOSAL SITES.   |          |     |        |
| CC | -1- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED.  |          |     |        |
| CC | -1- SUBCELLULAR LOCATION: Secreted.   |          |     |        |
| CC | -1- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN CYTOTOXIC T-CELLS.  |          |     |        |
| CC | -1- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.  |          |     |        |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ). |          |     |        |

#### ALIGNMENTS

|    |       |      |      |              |                    |
|----|-------|------|------|--------------|--------------------|
| 34 | 414.5 | 17.4 | 299  | 1 FCN3_HUMAN | O75636 homo sapien |
| 35 | 399   | 16.8 | 255  | 1 MFA4_HUMAN | P55083 homo sapien |
| 36 | 397.5 | 16.7 | 4289 | 1 TENX_HUMAN | P22105 homo sapien |
| 37 | 388   | 16.3 | 2201 | 1 TENA_HUMAN | P24821 homo sapien |
| 38 | 386.5 | 16.3 | 1746 | 1 TENA_PIG   | O29116 sus scrofa  |
| 39 | 386   | 16.2 | 1808 | 1 TENA_CHICK | P10039 gallus gall |
| 40 | 358.5 | 15.1 | 611  | 1 FIB2_PETMA | P33573 petromyzon  |
| 41 | 346.5 | 14.6 | 774  | 1 SCA_DROME  | P21520 drosophila  |
| 42 | 282   | 11.9 | 137  | 1 AGP2_RAT   | O35462 rattus norv |
| 43 | 185.5 | 7.8  | 129  | 1 MFA4_BOVIN | P55918 bos taurus  |
| 44 | 116.5 | 4.9  | 1005 | 1 RASQ_METJA | O58718 methanococc |
| 45 | 115   | 4.8  | 569  | 1 CYSF_PLAFA | P25805 plasmodium  |

```

DR EMBL: Z36531; CAA85298.1;
DR EMBL: AF104015; AAD10825.1;
DR EMBL: AF104014; AAD10825.1; JOINED.
DR EMBL: AF468959; AAL68855.1;
DR HSP: P02671; 1F2D.
DR GENE: HGNC:3696; FGL2.
DR MIM: 605351;
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; fibrinogen_C; 1.
DR SMART: SM00186; FBG; 1.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW T-cell; glycoprotein; signal; polymorphism.
FT SIGNAL 1 23
FT CHAIN 24 439
FT DOMAIN 210 435
FT DISULFID 213 242
FT DISULFID 371 384
FT CARBOHYD 25 25
FT CARBOHYD 179 179
FT CARBOHYD 235 235
FT CARBOHYD 263 263
FT CARBOHYD 336 336
FT VARIANT 53 53
FT SEQUENCE 439 AA; 50228 MW; DF34656288E49E68 CRC64;

Query Match 100.0%; Score 2378; DB 1; Length 439;
Best Local Similarity 100.0%; Pred. No. 2.8e-153;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLANWYSSAVLATYGLVANNETETIKDERAKDVCPRVLRSGKCEEGECPCYQVS 60
DB 1 MRLANWYSSAVLATYGLVANNETETIKDERAKDVCPRVLRSGKCEEGECPCYQVS 60
QY 61 LPPLTQLPKQSRLEEVKEVQNLKEIVNSLKKSCQDKLQADDNGDGRNGLLPSTG 120
DB 61 LPPLTQLPKQSRLEEVKEVQNLKEIVNSLKKSCQDKLQADDNGDGRNGLLPSTG 120
QY 121 APGEVGNVRRELESEVKNLSSELKNAKEEINVLHGRLEKLNVLNNNIENYVDSKVANL 180
DB 121 APGEVGNVRRELESEVKNLSSELKNAKEEINVLHGRLEKLNVLNNNIENYVDSKVANL 180
QY 181 TFVNSLDGKCKPQSOEQISQRPVQHLKYDCSDYVAGKRSSEYRTVTPDPKNSFEV 240
DB 181 TFVNSLDGKCKPQSOEQISQRPVQHLKYDCSDYVAGKRSSEYRTVTPDPKNSFEV 240
QY 241 YCDMETGGGWTVQLQRLDSTNFTRTWQDYKAGFGNLRREFWLGNDKTHLLTKSKEMIL 300
DB 241 YCDMETGGGWTVQLQRLDSTNFTRTWQDYKAGFGNLRREFWLGNDKTHLLTKSKEMIL 300
QY 301 RIDLEDFNGVELYALQDYVANEFLKYRLHGVNNGTAGDALRFNKHVNHDLKFFTPD 360
DB 301 RIDLEDFNGVELYALQDYVANEFLKYRLHGVNNGTAGDALRFNKHVNHDLKFFTPD 360
QY 361 KNDRTVPSNCGLYSSGWFDFACLSANLNGKYYHQYRGVRNGIFWGTWPGVSEAHPGG 420
DB 361 KNDRTVPSNCGLYSSGWFDFACLSANLNGKYYHQYRGVRNGIFWGTWPGVSEAHPGG 420
QY 421 YKSFKEAKMIRPKHKFP 439
DB 421 YKSFKEAKMIRPKHKFP 439

RESULT 2
FGL2_MOUSE
ID FGL2_MOUSE STANDARD; PRT; 432 AA.
AC P12804;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibroblast precursor (Fibrinogen-like protein 2) (Prothrombinase)
DE (Cytotoxic T-lymphocyte specific protein).
GN FGL2 OR FIBLP.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cytotoxic T-cell;
RX MEDLINE=87175527; PubMed=3550794;
RA Koyama T., Hall L.R., Hasegawa S., Saito H.;
RT "Structure of a cytotoxic T-lymphocyte-specific gene shows a strong
RT homology to fibrinogen beta and gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:1609-1613(1987).
RN [2]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC STRAIN=BALB/CJ; TISSUE=Peritoneal macrophage;
RX MEDLINE=9533285; PubMed=7690703;
RA Parr R.L., Fung L., Reneker J., Myers-Mason N., Leibowitz J.L.;
RA Levy G.;
RT "Association of mouse fibrinogen-like protein with murine hepatitis
RT virus-induced prothrombinase activity.";
RL J. Virol. 69:5033-5038(1995).
CC -!- FUNCTION: CONVERTS PROTHROMBIN TO THROMBIN.
CC -!- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN CYTOTOXIC
CC T-CELLS.
CC -!- INDUCTION: IN MACROPHAGES, DURING INFECTION BY MOUSE HEPATITIS
CC VIRUS STRAIN 3 (MHV-3).
CC -!- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M16238; AAA37624.1;
DR EMBL: M15761; AAA37624.1; JOINED.
DR EMBL: S78773; AAB34823.1;
DR PIR: A27447; A27447.
DR HSP: P02671; 1F2D.
DR MGD: MGI:103266; Fg12.
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; fibrinogen_C; 1.
DR SMART: SM00186; FBG; 1.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW T-cell; Cytolysis; Signal.
FT SIGNAL 1 19
FT CHAIN 20 432
FT DOMAIN 203 428
FT DISULFID 206 235
FT DISULFID 364 377
FT CARBOHYD 24 24
FT CARBOHYD 172 172
FT CARBOHYD 228 228
FT CARBOHYD 256 256
FT CARBOHYD 329 329
FT CARBOHYD 332 332
FT CONFLICT 332 332
FT SEQUENCE 432 AA; 48951 MW; 2B297F69CBB4A782 CRC64;

Query Match 77.9%; Score 1853.5; DB 1; Length 432;
Best Local Similarity 77.7%; Pred. No. 6.3e-118;
Matches 341; Conservative 42; Mismatches 49; Indels 7; Gaps 4;

QY 1 MRLANWYSSAVLATYGLVANNETETIKDERAKDVCPRVLRSGKCEEGECPCYQVS 60
DB 1 MRLANWYSSAVLATYGLVANNETETIKDERAKDVCPRVLRSGKCEEGECPCYQVS 60
QY 61 LPPLTQLPKQSRLEEVKEVQNLKEIVNSLKKSCQDKLQADDNGDGRNGLLPSTG 120
DB 61 LPPLTQLPKQSRLEEVKEVQNLKEIVNSLKKSCQDKLQADDNGDGRNGLLPSTG 120
QY 121 APGEVGNVRRELESEVKNLSSELKNAKEEINVLHGRLEKLNVLNNNIENYVDSKVANL 180
DB 121 APGEVGNVRRELESEVKNLSSELKNAKEEINVLHGRLEKLNVLNNNIENYVDSKVANL 180
QY 181 TFVNSLDGKCKPQSOEQISQRPVQHLKYDCSDYVAGKRSSEYRTVTPDPKNSFEV 240
DB 181 TFVNSLDGKCKPQSOEQISQRPVQHLKYDCSDYVAGKRSSEYRTVTPDPKNSFEV 240
QY 241 YCDMETGGGWTVQLQRLDSTNFTRTWQDYKAGFGNLRREFWLGNDKTHLLTKSKEMIL 300
DB 241 YCDMETGGGWTVQLQRLDSTNFTRTWQDYKAGFGNLRREFWLGNDKTHLLTKSKEMIL 300
QY 301 RIDLEDFNGVELYALQDYVANEFLKYRLHGVNNGTAGDALRFNKHVNHDLKFFTPD 360
DB 301 RIDLEDFNGVELYALQDYVANEFLKYRLHGVNNGTAGDALRFNKHVNHDLKFFTPD 360
QY 361 KNDRTVPSNCGLYSSGWFDFACLSANLNGKYYHQYRGVRNGIFWGTWPGVSEAHPGG 420
DB 361 KNDRTVPSNCGLYSSGWFDFACLSANLNGKYYHQYRGVRNGIFWGTWPGVSEAHPGG 420
QY 421 YKSFKEAKMIRPKHKFP 439
DB 421 YKSFKEAKMIRPKHKFP 439

```

|   |  |                        |
|---|--|------------------------|
| 121   | AFCEYGNRYRELESEVANKLSSELKNAKPEINVLHGRLEKLNLYNMNNIENYVDSKVANL         | 160                    |
| 122   | :     :     :     :     :     :     :     :     :     :              |                        |
| 123   | :     :     :     :     :     :     :     :     :                    |                        |
| 115   | AB-TAEDSRVOELESQVKNLSSELKNAKDOILOGQLGRLETLHLVNMNNIENYVDKNVANL        | 173                    |
| 116   | :     :     :     :     :     :     :     :     :                    |                        |
| 181   | TEVVASLIDGKCKSCPSQEOIQSRPVQHLIYKDCSDYVAIGKSSSETVRYTPDPKSSFEV         | 240                    |
| 182   | :     :     :     :     :     :     :     :     :                    |                        |
| 174   | TVVNSLIDGKCKSCPSQEHMQSPVQHLIYKDCSDHIVYIGRSSGAYRYTPDHRNSSFEV          | 233                    |
| 175   | :     :     :     :     :     :     :     :     :                    |                        |
| 241   | YCDMETMGGWTVLQARLDGSTNFTRWODYKAGGNLRERFWLGNDKHLHFKSEML               | 300                    |
| 242   | :     :     :     :     :     :     :     :     :                    |                        |
| 234   | YCDMETMGGWTVLQARLDGSTNFTREWKDYKAGFNLREFWLGNDKIHLLTKSEML              | 293                    |
| 235   | :     :     :     :     :     :     :     :     :                    |                        |
| 301   | RIDLEDFNGVELYALYDOFYVANEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFETPD         | 360                    |
| 302   | :     :     :     :     :     :     :     :     :                    |                        |
| 294   | RIDLEDFNGLTLALYDOFYVANEFLKYRLHIGNYNGTAGDALRESRIYNHDLRFFETPD          | 353                    |
| 295   | :     :     :     :     :     :     :     :     :                    |                        |
| 361   | KNDPRYSGNGGLYSSGWNFDACLSANLKYHYQKRYGVRNGIFWGTWPGVGSHEPGG             | 420                    |
| 362   | :     :     :     :     :     :     :     :     :                    |                        |
| 354   | RNDPRYSGNGGLYSSGWNFDCLSANLKYHYQKRYGVRNGIFWGTWPGINQAQFPG              | 413                    |
| 355   | :     :     :     :     :     :     :     :     :                    |                        |
| 421   | YKSSFEKAKMIRPKHFP 439  |                        |
| 422   | :     :     :     :     :     :     :     :     :                    |                        |
| 414   | YKSSFKQAKMIRPKNEP 432  |                        |
| 415   | :     :     :     :     :     :     :     :     :                    |                        |
| RESULT 3  |  |                        |
| AGP2_MOUSE  |  |                        |
| ID  | AGP2_MOUSE   | STANDARD; PRT; 496 AA. |
| IC  | O35608;  |                        |
| DD  | 16-OCT-2001 (Rel. 40, Created)                                       |                        |
| DT  | 16-OCT-2001 (Rel. 40, Last sequence update)                          |                        |
| DT  | 16-OCT-2001 (Rel. 40, Last annotation update)                        |                        |
| DE  | Angiopoietin-2 precursor (ANG-2).                                    |                        |
| GN  | Angiopoietin-2 precursor (ANG-2).                                    |                        |
| DE  | AGP2 OR AGPT2.   |                        |
| OS  | Mus musculus (Mouse).  |                        |
| OS  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |                        |
| OX  | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.   |                        |
| OX  | NCBI_TaxID=10090;  |                        |
| RP  | [1]  |                        |
| RN  | SEQUENCE FROM N.A.   |                        |
| RC  | TISSUE=Uterus;   |                        |
| RC  | MEDLINE=97349327; PubMed=9204896;                                    |                        |
| RA  | Maisonnier P.C., Suri C., Jones P.F., Bartunkova S., Wiegand S.J.,   |                        |
| RA  | Radotejewska C., Compton D.L., McClain J., Aldrich T.H.,             |                        |
| RA  | Papadopoulos N., Daly T.J., Davis S., Sato T.N., Yancopoulos G.D.;   |                        |
| RT  | "Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo |                        |
| RT  | angiogenesis."   |                        |
| RT  | Science 277:55-60(1997).   |                        |
| -1- FUNCTION: BINDS TO TIE2 RECEPTOR AND COUNTERACTS BLOOD VESSEL   |  |                        |
| MATURATION/STABILITY MEDIATED BY ANGIOPOIETIN-1. ITS FUNCTION MAY   |  |                        |
| BE CONTEXT-DEPENDENT. IN THE ABSENCE OF ANGIOGENIC INDUCERS, SUCH   |  |                        |
| AS VEGF, ANG2-MEDIATED LOOSENING OF CELL-MATRIX CONTACTS MAY  |  |                        |
| INDUCE ENDOTHELIAL CELL APOPTOSIS WITH CONSEQUENT VASCULAR  |  |                        |
| REGRESSION. IN CONCERT WITH VEGF, IT MAY FACILITATE ENDOTHELIAL   |  |                        |
| CELL MIGRATION AND PROLIFERATION, THUS SERVING AS A PERMISSIVE  |  |                        |
| ANGIOGENIC SIGNAL.  |  |                        |
| -1- SUBCELLULAR LOCATION: Secreted.   |  |                        |
| -1- TISSUE SPECIFICITY: EXPRESSED ONLY AT SITES OF VASCULAR   |  |                        |
| REMODELING.   |  |                        |
| -1- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.  |  |                        |
| CC  |  |                        |
| CC This SWISS-PROT entry is copyright. It is produced through a collaboration   |  |                        |
| CC between the Swiss Institute of Bioinformatics and the EMBL outstation  |  |                        |
| CC the European Bioinformatics Institute. There are no restrictions on its  |  |                        |
| CC use by non-profit institutions as long as its content is in no way   |  |                        |
| CC modified and this statement is not removed. Usage by and for commercial  |  |                        |
| CC entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> |  |                        |
| CC or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).  |  |                        |
| CC  |  |                        |
| CC EMBL; AF004326; AAB63189.1; -  |  |                        |
| CC DR HSSP; P02671; 1FZD.   |  |                        |
| CC DR MGD; MGI:1202890; Agpt2.  |  |                        |



```

FT DOMAIN 280 496 FIBRINOGEN C-TERMINAL.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 97 148 MISSING (IN ISOFORM 2).
SQ SEQUENCE 496 AA; 56919 MW; 5642A58847A7385C CRC64;

Query Match 22.0%; Score 522; DB 1; Length 496;
Best Local Similarity 30.4%; Pred. No. 4.3e-28;
Matches 136; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE--EAGECPYOVSLPPL-----TIQ--LPKQFSRIEVEKVEON 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 MDSIGKKQYOVGSCSYFLPDMDCRSSSPYVNAVORDAPLEYDDSVORLOVLEN 85
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 85 LKE-----IVNSLKKSCQCKLOADDNGPGRNGLLLPSTGAPGEVGN----- 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
86 IMENNTQWLKLENYIQDNMKKEMVEIQQNAVON-----QTAVMIETGNTLLNQ 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 -----RVRELESEV-----NKLSELKNKEEINVLHGR---LEKL 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 TAEQTRKLTVEAQLVNTQTRLEQLLHLSLSTNKLKQLDQTSLEINKLQDNKSFLEKK 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 162 NLV-----NMNIENYVDSKVANLTF-----V 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 VLAMEDKHIIQLOSKEERDQLOVLVSKNSIELEKKIYVATVNSVLQKQOHDLMET 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 VNSLDGKSCPCSQEQIOSRPVQHLIYKDCSDYAIAGKRSSERYVPDPKNSSEVYCD 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 VNNLLTWMSTNSAKDPTVAKEEQISFDCAEYKSGHTNGIYTLFPNSTEIKAYCD 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 MTMGSGWTVLQARLDGNTNTRWQDYKAGFNLGRFELWGNKDIHLLTKKEMILRID 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 MEAGSGWTIQRREDSYDFQRTWKYKVGFGNPSGSEYVWLNFEVSQLTNQORVYLIH 374
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 304 LEDEFNGVELYALDYQYVANEFLKYLHVGNVNGTAGDALRNFKNHDLKFTTDPKDN 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 375 LKDWEGNEAYSLYEHFLSSELNRYIHLKGLTAGKISSIQPND---FSTKQDGN 430
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 364 DRYPSGNGCLYSGGWFDACLSANLNGKYYHOKYRGVYR-NGIPIGWTPVSEAHPGYK 422
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 431 DKRCIC-KCSQMLTGGWFDACGPNLNGMYIPQONTNFKNGIKWYWKG-----SGY- 482
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 423 SSFKEAKNMIRPKHE 437
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
483 -SLKATTIMIRPADF 496

RESULT 6
FIBB_CHICK
ID FIBB_CHICK STANDARD; PRT; 463 AA.
AC Q02020;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibrinogen beta chain precursor [Contains: Fibrinopeptide B]
DE (Fragment).
GS F0B.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-13 AND 18-39.
RX MEDLINE=91182745; PubMed=2009266;
RA Weissbach L., Oddoux C., Procyk R., Grieninger G.;
RT "The beta chain of chicken fibrinogen contains an atypical thrombin
RT cleavage site."
RL Biochemistry 30:3290-3294 (1991).

```

```

CC CC FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC CC AGGREGATION.
CC CC SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC CC MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT. THE SOFT CLOT IS
CC CC CONVERTED INTO THE HARD CLOT BY FACTOR XIII WHICH CATALYZES THE
CC CC ESTERON-(GAMMA-GLUTAMYL)LYSINE CROSS-LINKING BETWEEN GAMMA CHAINS
CC CC (STRONGER) AND BETWEEN ALPHA CHAINS (WEAKER) OF DIFFERENT
CC CC MONOMERS.
CC CC SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: M58514; AAA48770.1; .
DR PIR: A38463; A38463.
DR HSSP: P02675; 1F2F.
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; fibrinogen_C; 1.
DR SMART: SM00186; FBG; 1.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Blood coagulation; Plasma; Platelet; Glycoprotein; Sulfation.
FT NON_TER 1 1
FT PEPTIDE <1 17 FIBRINOPEPTIDE B.
FT CHAIN 18 463 FIBRINOGEN BETA CHAIN.
FT MOD_RES 5 5 SULFATION (BY SIMILARITY).
FT SITE 17 18 CLEAVAGE (BY THROMBIN; RELEASE
FT FIBRINOPEPTIDE B).
FT DISULFID 69 69 INTERCHAIN (WITH ALPHA) (BY SIMILARITY).
FT DISULFID 80 80 INTERCHAIN (WITH ALPHA) (BY SIMILARITY).
FT DISULFID 84 84 INTERCHAIN (WITH GAMMA) (BY SIMILARITY).
FT DISULFID 197 197 INTERCHAIN (WITH ALPHA) (BY SIMILARITY).
FT DISULFID 201 201 INTERCHAIN (WITH GAMMA) (BY SIMILARITY).
FT DISULFID 205 289 BY SIMILARITY.
FT DISULFID 215 244 BY SIMILARITY.
FT DISULFID 397 410 BY SIMILARITY.
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 463 AA; 52678 MW; 2044CD49BA79EC7B CRC64;

Query Match 21.5%; Score 511.5; DB 1; Length 463;
Best Local Similarity 30.4%; Pred. No. 2e-27;
Matches 137; Conservative 71; Mismatches 154; Indels 89; Gaps 16;

QY 31 KDERAKDVCVRLESRGK-----BEAGE-CPYQVSLPPLITTLQPKQFSRIEVEKVEON- 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 QDKQAMKKGPIIYPDAGCKHPLDELGLVLCPTGCE---LQTTLLKQKTVKPYLRDLKDR 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 85 -----LKEIVNSLKKSCQCKLQADDNGPGRNGLLLPSTGAPGEVGNVRELE 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 YAKFSDSTTMYQYVNMIDNKLVTQKQKRD-----NDIILSEYNTMELHYNVIK--D 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 SEVKNLSSELKNKEEINVLHGRLEKLNLYNMNIENYVDSKVANLTFVNSLDGKSCK 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 NLDNIFSSRLVRUAVYDLSLHKIKOKL-----ENAIATQ-----DYCRGPC 201
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 195 PSEOIQOSRPVQHLIYKDCSDYAIAGKRSSERYVPDPKNSSEVYCDMETGGWTVL 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 202 -----VASCNIPVSGRECEDIYKGGTSETMYIOPDFTTPYRVYCDMETDNGWTLI 256
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 255 QARLDGNTNTRWQDYKAGFNLGR-----EFWLGNDKIHLTKKEMILRID 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 QNRQDGSVNGRAWDEYKRGFGNIAGSGKKGYCDTPGEYLGNDKIKSLQTLIGPTKVLI 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 304 LEDEFNGVELYALDYQYVANEFLKYLHVGNVNGTAGDALR--FNKHYN-----HDLK 354

```

```

Db 317 MEDWNGDKVSALYGGFTIHNEGKNYQLSVSNYKGNAGNALMEGASQLYGENRTWTIINGM 376
QY 355 FFTTPDKDNDRY----PSGNGCLYSSGWMFADCLSANLNGKY-----HQYRGVNRN 403
Db 377 YFSTYDRDNDGWTLPDKKCKSDGGGWNYNRCHAANPNRGYYGGTYSNDMAKHGTD 436
QY 404 GIEWGTWPGVSEAHGPGGYSFKEAKMMIRP 434
Db 437 GIVMNMWKG-----SWY--SMKKMSMKIKP 459

RESULT 7
FGLL_HUMAN
ID FGLL_HUMAN STANDARD; PRT; 312 AA.
AC Q08830; Q96Q6; Q96K6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen-like protein 1 precursor (Hepatocyte-derived fibrinogen-
related protein 1) (HFEPI-1) (Hepassocin) (HP-041).
GN FGLI OR HFEPI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93290661; PubMed=8390249;
RA Yamamoto T., Gotoh M., Sasaki H., Terada M., Kitajima M.,
RA Hirohashi S.;
RT "Molecular cloning and initial characterization of a novel
RT fibrinogen-related gene, HFEPI-1."
RL Biochem. Biophys. Res. Commun. 193:681-687(1993).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=21363035; PubMed=11470158;
RA Hara H., Yoshimura H., Uchida S., Toyoda Y., Aoki M., Sakai Y.,
RA Morimoto S., Shikawa K.;
RT "Molecular cloning and functional expression analysis of a cDNA for
RT human hepassocin, a liver-specific protein with hepatocyte mitogenic
RT activity."
RL Biochim. Biophys. Acta 1520:45-53(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX Strausberg R.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Has hepatocyte mitogenic activity.
CC -!- SUBUNIT: Homodimer (Probable).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Liver-specific.
CC -!- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14446; BAA03336.1;
DR EMBL: D87342; BAB70690.1;
DR EMBL: BC007047; AA07047.1;
DR HSSP: P02671; 1FZD.
DR Genew: HGNC:3695; FGLI.
DR Ref: 605776;
DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; Fibrinogen_C; 1.
DR SMART: SM00186; FBG; 1.

```

```

DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Signal.
FT SIGNAL 1 22
FT CHAIN 23 312 FIBRINOGEN-LIKE PROTEIN 1.
FT DOMAIN 78 305 FIBRINOGEN C-TERMINAL.
FT DISULFID 26 26 INTERCHAIN (POTENTIAL).
FT DISULFID 83 112 BY SIMILARITY.
FT DISULFID 248 261 BY SIMILARITY.
FT CONFLICT 15 15 I -> T (IN REF. 3).
FT CONFLICT 69 69 N -> D (IN REF. 1).
FT CONFLICT 72 72 I -> V (IN REF. 1).
FT CONFLICT 105 105 P -> L (IN REF. 2).
SQ SEQUENCE 312 AA; 36391 MW; 268C82124E6660C2 CRC64;

Query Match 21.3%; Score 506.5; DB 1; Length 312;
Best Local Similarity 36.8%; Pred. No. 2.7e-27;
Matches 119; Conservative 40; Mismatches 119; Indels 45; Gaps 10;

QY 131 RELESEVKNLSSELKNAKEINVHGRLE----KLNLYNMNNIENYVDSKVANLTFVYNS 186
Db 18 RET-SALEDCAQEQMRLRAQVRLLETRVKKQOVKIKQLQENEVQFLDKGDENVIDLG- 75
QY 187 LDGKCKSPQEQIQSRPVQHLIYKDCSDYVAGKRSSTYRTVTPDKNSSFEVYCDMET 246
Db 76 -----SKRQ-----YADCEIENDYKLSGFYKIKPLQSPAFESVYCDMSD 116
QY 247 MGGGWTVIQARLDGSTNFTTWODYKAGFNL---RREFWLGNDKIHLTKSKEMILRID 303
Db 117 -GGGWTVIQRRSDGSENFNRGKWKYENGFGFKHGEYWLGNKLNHLFTTQEDYTLKID 175
QY 304 LEDFNGVELYALYDQFYVANEFLKYLHVNGNYNTAGDALRENFH-----YNHDLKFEFT 357
Db 176 LADFEKNSRYAQYKFKVGDENFYELNIGESTAGDSLACNFHPEVQWVWASHQRMKES 235
QY 358 TPDKNDRYPGNGCLYSSGWMFADCLSANLNGKYHOKYRG-VRNGIFWGTWPGVSEA 416
Db 236 TWRDHDNY-EGNCAEDQSGWFWNRCHSANLNGVYSGPYTAKTDNGIVMYTW----- 288
QY 417 HPGGYKSSFSFKEAKMIRPKHFKP 439
Db 289 --HGWWYSLKSVYMKIRPNDFIP 309

RESULT 8
AGPI_HUMAN
ID AGPI_HUMAN STANDARD; PRT; 498 AA.
AC Q15389;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Angiotensin-1 precursor (ANG-1).
GN ANGPT1 OR KIAA0003.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT GLY-269 DEL.
RC TISSUE=Petal lung;
RX MEDLINE=97134663; PubMed=8980223;
RA Davis S., Aldrich T.H., Jones P.F., Acheson A., Compton D.L., Jain V.,
RA Ryan T.E., Bruno J., Radziejewski C., Maissonpiere P.C.,
RA Vancopoulos G.D.;
RT "Isolation of angiotensin-1, a ligand for the TIE2 receptor, by
RT secretion-trap expression cloning."
RL Cell 87:1161-1169(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Ohara O., Nagase T., Kikuno R., Nomura N.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 307-498 FROM N.A.
RC TISSUE=Bone marrow;

```





RA Farrel D.H., Mulvihill E.R., Huang S., Chung D.W., Davie E.W.;  
 RL "Recombinant human fibrinogen and sulfation of the gamma chain."; [16]  
 RN Biochemistry 30:9414-9420(1991).  
 RP REVIEW, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.  
 RX MEDLINE=84305751; PubMed=6383194;  
 RA Doolittle R.F.;  
 RL "Fibrinogen and fibrin."; [17]  
 RN Annu. Rev. Biochem. 53:195-229(1984).  
 RP POLYMERIZATION SITE.  
 RX MEDLINE=85014892; PubMed=6592597;  
 RA Horwitz B.H., Varadi A., Scheraga H.A.;  
 RL "Localization of a fibrin gamma-chain polymerization site within segment Thr-374 to Glu-396 of human fibrinogen."; [18]  
 RN Proc. Natl. Acad. Sci. U.S.A. 81:5980-5984(1984).  
 RP POLYMERIZATION SITE.  
 RX MEDLINE=81142375; PubMed=6451630;  
 RA Olexa S.A., Budzynski A.Z.;  
 RL "Localization of a fibrin polymerization site."; [19]  
 RN J. Biol. Chem. 256:3544-3549(1981).  
 RP PLATELET AGGREGATION SITE.  
 RX MEDLINE=84203545; PubMed=6326808;  
 RA Kloczewiak M., Timmons S., Lukas T.J., Hawiger J.;  
 RL "Platelet receptor recognition site on human fibrinogen. Synthesis and structure-function relationship of peptides corresponding to the carboxy-terminal segment of the gamma chain."; [20]  
 RN Biochemistry 23:1767-1774(1984).  
 RP PLATELET AGGREGATION SITE.  
 RX MEDLINE=84185664; PubMed=6325435;  
 RA Ploew E.F., Strouji A.H., Meyer D., Marquerie G., Ginsberg M.H.;  
 RL "Evidence that three adhesive proteins interact with a common recognition site on activated platelets."; [21]  
 RN J. Biol. Chem. 259:5388-5391(1984).  
 RP CALCIUM-BINDING SITE.  
 RX MEDLINE=85261382; PubMed=3160702;  
 RA Dang C.V., Ebert R.F., Bell W.R.;  
 RL "Localization of a fibrinogen calcium binding site between gamma-subunit positions 311 and 336 by terbium fluorescence."; [22]  
 RN J. Biol. Chem. 260:9713-9719(1985).  
 RP CHROMATOGRAPHIC COMPARISON OF GAMMA-A AND GAMMA-B CHAINS.  
 RX MEDLINE=81054908; PubMed=6933547;  
 RA Wolfenstein-Todel C., Mosesson M.W.;  
 RL "Human plasma fibrinogen heterogeneity: evidence for an extended carboxyl-terminal sequence in a normal gamma chain variant (gamma')."; [23]  
 RN Proc. Natl. Acad. Sci. U.S.A. 77:5069-5073(1980).  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 169-437.  
 RX MEDLINE=97169449; PubMed=9016719;  
 RA Yee V.C., Pratt K.P., Cote H.C.F., le Trong I., Chung D.W., Davie E.W., Stenkamp R.E., Teller D.C.;  
 RL "Crystal structure of a 30 kDa C-terminal fragment from the gamma chain of human fibrinogen"; Structure 5:125-138(1997). [24]  
 RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 169-437.  
 RX MEDLINE=97352771; PubMed=9207064;  
 RA Pratt K.P., Cote H.C.F., Chung D.W., Stenkamp R.E., Davie E.W.;  
 RL "The primary fibrin polymerization pocket: three-dimensional structure of a 30-kDa C-terminal gamma chain fragment complexed with the peptide Gly-Pro-Arg-Pro."; [25]  
 RN Proc. Natl. Acad. Sci. U.S.A. 94:7176-7181(1997).  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 114-432.  
 RX MEDLINE=97472408; PubMed=9333233;  
 RA Spraggon G., Everse S.J., Doolittle R.F.;  
 RL "Crystal structures of fragment D from human fibrinogen and its

RT crosslinked counterpart from fibrin."; [26]  
 RN Nature 389:455-462(1997).  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 114-432.  
 RX MEDLINE=98292395; PubMed=9628725;  
 RA Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;  
 RL "Crystal structure of fragment double-D from human fibrin with two different bound ligands."; [27]  
 RN Biochemistry 37:8637-8642(1998).  
 RP X-RAY CRYSTALLOGRAPHY.  
 RX MEDLINE=99175089; PubMed=10074346;  
 RA Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;  
 RL "Conformational changes in fragments D and double-D from human fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide."; [28]  
 RN Query Match 20.4%; Score 484.5; DB 1; Length 453;  
 Best Local Similarity 29.7%; Pred. No. 1.3e-25;  
 Matches 141; Conservative 56; Mismatches 178; Indels 99; Gaps 18;  
 QY 8 WLSSAVLATYGFVANNTEETIKDERAKDVCVRLESRGKCEEAGECPYQVSLPPLTIQ 67  
 DB 19 FLSSTCVA-----YVATRDNCCILDERFGSYCPT-----TCGIADFLSTYQTK 61  
 QY 68 LKQFSRTEEVFKEVQNLEIVNSLKKSCQCKLQADDNGDGRNGLLLPSTGAPGVD 127  
 DB 62 VDKDQSLDELHQVENKTSYEVKQLIKAIQ-LTYPNDESKPNMT-----DAAT 109  
 QY 128 NRVRESEVANKLSSELKNAKEEINVLHGRLEKLNVMNNTENYVDSKVANLTFVNSL 187  
 DB 110 LKSRKMLEEIMKYEASILTHDSSIRYLO-----EYNSNN-----QKIVNLKKEVAOL 157  
 QY 188 DKCKSCQSQOIQSRVQHLIYKDCSDYTAIGRSRSEYRTPDPKNSFEYCDMETM 247  
 DB 158 EAQCE-PCKDTVO---IHDITGDKCDQIANKAKQSGLYFKPLKANQOQFLYCYBIDS 213  
 QY 248 GGGVTVLQARLDGSTNFTRTQDYKAGFNL-----RREFWLGNDKIHLTKSKEM--ILR 301  
 DB 214 GNGVTVECKRLDGSVDKKNWQIQYKEGHLSPGTTFEFLGNEKIHILISTQSAIYALR 273  
 QY 302 IDLEDFNGVELYALYQFVYANFLKYLHVGNY-NGTAGDAL-RFNKHYNHLKFTT-- 357  
 DB 274 VELEDWNGRTSTADYAMFKVGPEDKYLTYVYFAGGDAGDAFGDFGDDPSDKFTTSH 333  
 QY 358 -----TPDKNDNRYPSGNGLYYSSGWFEDACLNSLNGKYYH-----QKYRGVRNG 404  
 DB 334 NGMQSTWDNDNKF-EGNCABQDQSGWMMNKCHAGLNGVYQGYYSKASTPNGYDNG 392  
 QY 405 IFWGTWP-----GVSEA---HPGKYKSFREKAKMIRPKH 436  
 DB 393 IIVATWKTWYSMKTKMTKIIPNRLTIGEGQQHILGGAK-----QVRPEH 438  
 RESULT 11  
 AGPL MOUSE STANDARD; PRT; 498 AA.  
 ID AGPL MOUSE STANDARD; PRT; 498 AA.  
 AC 008538;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Angiopoietin-1 precursor (ANG-1).  
 GN ANGPT1 OR AGPT.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID:10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97134663; PubMed=8980223;  
 RA Davis S., Aldrich T.H., Jones P.F., Acheson A., Compton D.L., Jain V., Ryan T.E., Bruno J., Radziejewski C., Maisonnier P.C., Yancopoulos G.D.;  
 RL "Isolation of angiopoietin-1, a ligand for the Tie2 receptor, by

RT secretion-trap expression cloning.";  
RL Cell 87:1161-1169(1996).  
CC -!- FUNCTION: BINDS AND ACTIVATES TIE2 RECEPTOR BY INDUCING ITS  
CC TYROSINE PHOSPHORYLATION. IMPLICATED IN ENDOTHELIAL DEVELOPMENTAL  
CC PROCESSES LATER AND DISTINCT FROM THAT OF VEGF. APPEARS TO PLAY A  
CC CRUCIAL ROLE IN MEDIATING RECIPROCAL INTERACTIONS BETWEEN THE  
CC ENDOTHELIUM AND SURROUNDING MATRIX AND MESENCHYME. MEDIATES BLOOD  
CC VESSEL MATURATION/STABILITY. IT MAY PLAY AN IMPORTANT ROLE IN THE  
CC HEART EARLY DEVELOPMENT.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DEVELOPMENTAL STAGE: EARLY IN DEVELOPMENT, AT E9 TO E11, IT IS  
CC FOUND MOST PROMINENTLY IN THE HEART MYOCARDIUM SURROUNDING THE  
CC ENDOCARDIUM. LATER, IT BECOMES MORE WIDELY DISTRIBUTED, MOST OFTEN  
CC IN THE MESENCHYME SURROUNDING DEVELOPING VESSELS, IN CLOSE  
CC ASSOCIATION WITH ENDOTHELIAL CELLS.  
CC -!- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U83509; AAB50558.1; -;  
DR HSSP: P02671; 1FZD.  
DR MGD: MGI:108448; Agpt.  
DR InterPro: IPR002181; Fibrinogen\_C.  
DR Pfam: PF00147; fibrinogen\_C; 1.  
DR SMART: SM00186; FBG; 1.  
DR PROSITE: PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
KW Glycoprotein; Coiled coil; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 498  
FT DOMAIN 81 119  
FT DOMAIN 153 261  
FT DOMAIN 284 498  
FT CARBOHYD 92 92  
FT CARBOHYD 122 122  
FT CARBOHYD 154 154  
FT CARBOHYD 243 243  
FT CARBOHYD 295 295  
FT CARBOHYD 498 498  
FT SEQUENCE 498 AA; 57505 MW; 285B4FDEC260D800 CRC64;  
-----  
Query Match 20.3%; Score 483.5; DB 1; Length 498;  
Best Local Similarity 31.7%; Pred. No. 1.7e-25;  
Matches 127; Conservative 62; Mismatches 141; Indels 71; Gaps 14;  
-----  
QY 61 LPPLTIOLPKQFSRIE-----EVEFKEVONLKEITVNSLAKSCQDCKLQADDNGDPCRN 112  
DB 145 LTVDTQVLNQTSLRIQLLENSLSTYKLEKQLLOQTNEILKHE-----KN 191  
QY 113 GILLPSTGAPGVGNRVRLESEVKNLSSELKNAKEEINVLHGR-----LEKLNVLN 165  
DB 192 SIL-----EHLKLEMEGKHKEEDLTKEENKLGVSQRQFTIQELEX-QLSR 239  
QY 166 MNINENYVDSKVANLTFVNSLDGCKSK---CPQBOIQSRPVQHLVYKQCSDYATGK 221  
DB 240 ATNNNSILQKQLEMDTVHNLVSLCTKRGVLLKGGKREBK-----FDCADVYQAGF 294  
QY 222 RSSEYTRV----TPDFKNSFEVYCDMETMGGGTVLQARLDGSTNFTRTWQYKAGFGN 277  
DB 295 NKSGYITTYTNNMPEPK----KVFCDNDVNGGWTVIQHREDGSLDPQRKWEYKMGFGN 350  
QY 278 LRREPLGNDKIHLTKSKEMLRDLIDEPNGVELYALDYQFYVANEFLKYLRLHGVNGY 337  
DB 351 PSGEVWLGNEFTAITTSQRQWLRIELMDWEGNRAYSQYDFHIGNEKQYANLYLKGTG 410  
QY 338 TAGDALRFNKHYNHDLKFTFTTPDKDNDYPSGNCGLYYSSGWPFDACLSNALNGKYHQ- 396  
DB 411 TAGKQSSILLH-GAD---FSTKQADNDNCMC-KCALMLTGWGFDACGFSNLNGMFTAG 465

QY 397 KYRGVRNGLFNCTWPGVSEAHPGGYKSKSEKAKMMIREKHF 437  
DB 466 QNHGKLGNGIKWHYFKGPS-----YSLRSTTMMIRPLDF 498  
-----  
RESULT 12  
ANL2\_HUMAN  
ID ANL2\_HUMAN STANDARD; PRT: 493 AA.  
AC O9UKU9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Angiopoietin-related protein 2 precursor (Angiopoietin-like 2):  
GN ANGPTL2 OR ARP2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Heart;  
RX MEDLINE=99403103; PubMed=10473614;  
RA Kim I., Moon S.-O., Koh K.N., Kim H., Uhm C.-S., Kwak H.J., Kim N.-G.,  
RA Koh G.Y.;  
RT "Molecular cloning, expression, and characterization of angiopoietin-  
RT related protein. angiopoietin-related protein induces endothelial cell  
RT sprouting.";  
RL J. Biol. Chem. 274:26523-26528(1999).  
CC -!- FUNCTION: INDUCES SPROUTING IN ENDOTHELIAL CELLS THROUGH AN  
CC AUTOCRINE AND PARACRINE ACTION.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN HEART, SMALL INTESTINE,  
CC SPLEEN AND STOMACH. ALSO FOUND IN LOWER LEVELS IN COLON, OVARY,  
CC ADRENAL GLAND, SKELETAL MUSCLE AND IN PROSTATE.  
CC -!- PTM: N-GLYCOSYLATED.  
CC -!- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF125175; AAD55357.1; -;  
DR HSSP: P02671; 1FZD.  
DR Genew: HGNC:490; ANGPTL2.  
DR MIM: 605001; -;  
DR InterPro: IPR002181; Fibrinogen\_C.  
DR Pfam: PF00147; fibrinogen\_C; 1.  
DR SMART: SM00186; FBG; 1.  
DR PROSITE: PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
KW Signal; Coiled coil; Glycoprotein.  
FT SIGNAL 1 22  
FT CHAIN 23 493  
FT DOMAIN 76 115  
FT DOMAIN 152 206  
FT DOMAIN 438 450  
FT CARBOHYD 164 164  
FT CARBOHYD 192 192  
FT SEQUENCE 493 AA; 57104 MW; 0F2ADECE53D185CA CRC64;  
-----  
Query Match 20.2%; Score 480.5; DB 1; Length 493;  
Best Local Similarity 30.8%; Pred. No. 2.7e-25;  
Matches 117; Conservative 62; Mismatches 128; Indels 73; Gaps 12;  
-----  
QY 66 IQLPKQFSRIEVEFVEV-----QNLKEIVNSLAKSCQDCKLQADDNGDPCRNGLLLPSTGA 121  
DB 170 LQASKYKQLEHKYOHATLTAHNOSEITIAQLEEHQCV-----PSAREVPOPPFAA 220  
QY 122 PGEVGD-----NRVRESEVSEVKNLSSELKNAKEEINVLHGRLEKLNVLNMMNTNIYVDSK 177





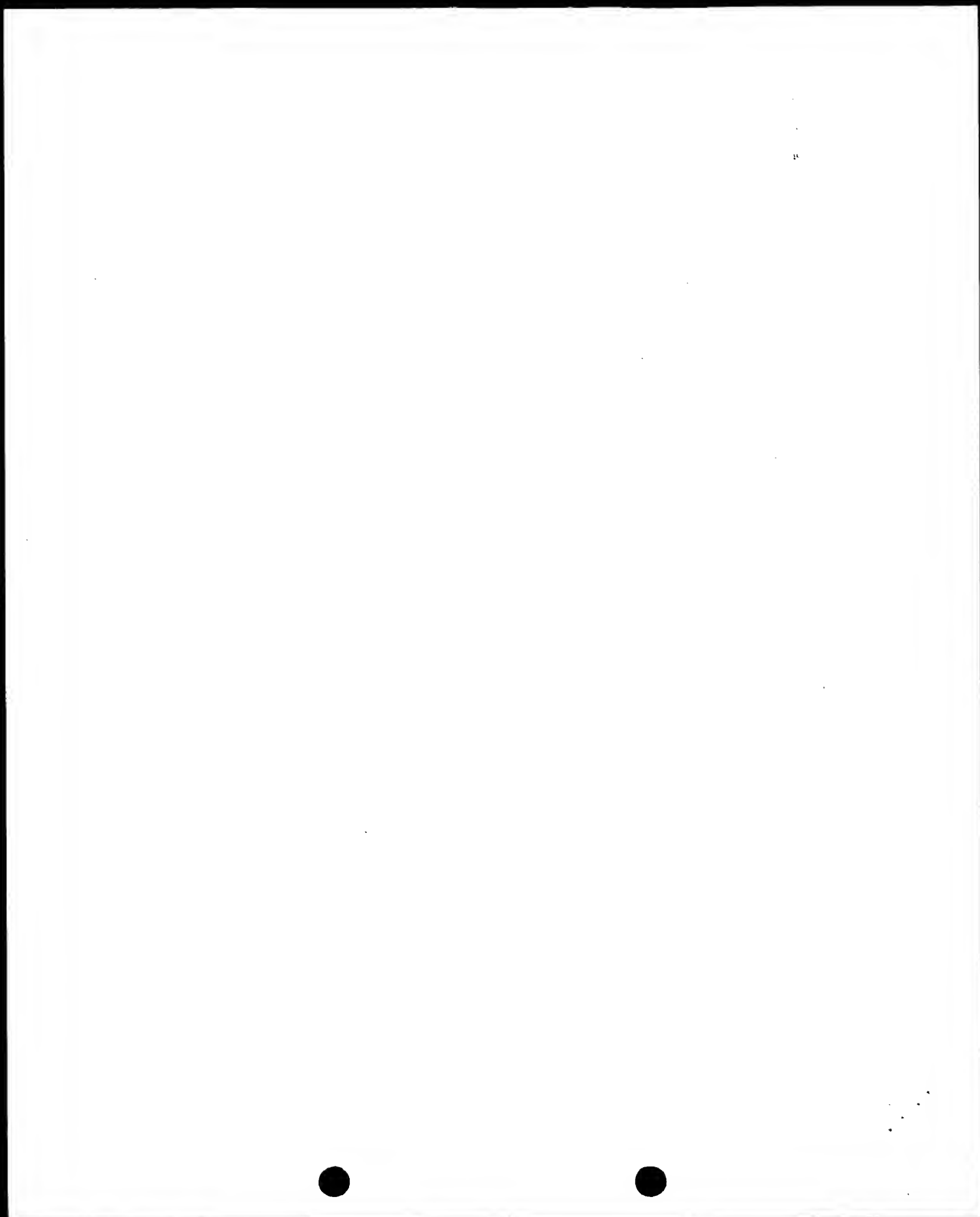
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; AF113707; AAD21586.1; -.  
DR HSP; P02671; 1FZD.  
DR MGD; MGI:1336887; Agpt4.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR Pfam; PF00147; fibrinogen\_C; 1.  
DR SMART; SM00186; FBG; 1.  
DR PROSITE; PS00314; FIBRIN\_AG\_C\_DOMAIN; 1.  
KW Coiled coil; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 509 ANGIOPOIETIN-4.  
FT DOMAIN 181 269 COILED COIL (POTENTIAL).  
FT DOMAIN 294 480 FIBRINOGEN-LIKE.  
FT SEQUENCE 509 AA; 57805 MW; 985A74A20A6664F6 CRC64;  
-----

Query Match 19.8%; Score 470.5; DB 1; Length 509;  
Best Local Similarity 29.0%; Pred. No. 1.3e-24;  
Matches 125; Conservative 74; Mismatches 154; Indels 78; Gaps 13;

QY 67 QLPKQFSRTVEEFK-EVQNLKEIVNSLKKSCQCKLOADDNGDPCGRNGLLLPSTGA----- 121  
DB 91 QRAQRYSOLEKLENNQWLLKLEQSIKVNLRSHLVQAQODITIQOTTML-ALGANLMN 149  
QY 122 PGEVGNRVRELESEV-----NKLSELKNAKEEINVLHGR----- 157  
DB 150 QTKAQTHKLTAVEAQLVQNLTKMLQLENSLSTNKLQRLQGRNRALET 209  
QY 158 ----LEKLVNMNIE-----NYVDSKVANLTFVNSLDGCKSCPCSOEQIQSR 203  
DB 210 RLQALEAQHQALNSLOEKREQLHSLGHQTCTLANLKHNLHLSNSSLSQQQQOOLTE 269  
QY 204 PVQHL-----IYKDCSDYYAIGKRSEYTVPTDPKNSFEVYCDME 245  
DB 270 FVQRLVRIVAQDHPVSLKTPKPFQDCAEIKRSCVNTSGVYIYETNMTKPKVFCDE 329  
QY 246 TMGGGWTVLQALDQSTNFTRTWODYKAGFGNLRREFWLGNQKIHLLTKSKEMILRLDLE 305  
DB 330 TDGGGWTLIOHREDGSVNFORTWEEYKEGFGNVAHEHVLGNVHRLTSRTAYLLRVELH 389  
QY 306 DFNGVELYALDQFYVANEFLKYRLHVGNYNCTAGDALRFNKHYNHDLKFFTTDPKDNDR 365  
DB 390 DWEGROTSTOVENFQLGSRQRYSLSVNDSSSSAG---RKNSLAPQGTG-FSTKMDNDN 445  
QY 366 YPSGNGCLYSSGWMFDACLSANLNGKYY--HOKYRGVRNGIFWGTWPCVSEAHPCGYKS 423  
DB 446 CMC-KCAQMLSGWNFDACLSNLNGCIYYSVHQHLHI-NGIRWYIFRGPSYSLHG----- 499  
QY 424 SFREAKMMIRP 434  
DB 500 ----TRMMLRP 506

Search completed: December 2, 2002, 06:51:11  
Job time : 38.7445 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:48:20 ; Search time 119.903 Seconds  
(without alignments)  
754.398 Million cell updates/sec

Title: US-09-902-563-2

Perfect score: 2378

Sequence: 1 MKLANNYWLSSAVLATYGFLL.....GYKSSFKKAKMIRPKHKP 439

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 1647  | 69.3        | 357    | 11    | Q9EPT7      |
| 2          | 750   | 31.5        | 148    | 4     | Q8WWE4      |
| 3          | 618.5 | 26.0        | 652    | 5     | Q5ND01      |
| 4          | 593.5 | 25.0        | 220    | 5     | Q8T8A2      |
| 5          | 526   | 22.1        | 496    | 11    | Q9D2D2      |
| 6          | 524   | 22.0        | 407    | 13    | Q9DER1      |
| 7          | 524   | 22.0        | 441    | 13    | Q9DER0      |
| 8          | 524   | 22.0        | 493    | 13    | Q9DER2      |
| 9          | 523.5 | 22.0        | 495    | 4     | Q9P2Y7      |
| 10         | 523   | 22.0        | 407    | 13    | Q9PU54      |
| 11         | 523   | 22.0        | 496    | 6     | Q8BDY7      |
| 12         | 518.5 | 21.8        | 314    | 11    | Q8VC25      |
| 13         | 505   | 21.2        | 491    | 4     | Q95841      |
| 14         | 496   | 20.9        | 244    | 4     | Q9HBP3      |
| 15         | 491.5 | 20.7        | 513    | 13    | Q90Z19      |
| 16         | 491   | 20.6        | 337    | 11    | Q8RLQ3      |

|    |       |      |      |    |        |
|----|-------|------|------|----|--------|
| 17 | 489   | 20.6 | 346  | 4  | Q43827 |
| 18 | 485.5 | 20.4 | 292  | 5  | Q9U8W8 |
| 19 | 482.5 | 20.3 | 488  | 13 | Q91589 |
| 20 | 481.5 | 20.2 | 493  | 11 | Q9JJ03 |
| 21 | 476   | 20.0 | 308  | 5  | Q9U8W6 |
| 22 | 475   | 20.0 | 435  | 13 | Q93568 |
| 23 | 474.5 | 20.0 | 498  | 6  | Q9BDY8 |
| 24 | 473   | 19.9 | 489  | 13 | Q90Z18 |
| 25 | 471   | 19.8 | 316  | 5  | Q9U8W7 |
| 26 | 456.5 | 19.2 | 324  | 5  | Q95P99 |
| 27 | 456   | 19.2 | 341  | 5  | Q966W1 |
| 28 | 454.5 | 19.1 | 431  | 6  | Q95LU3 |
| 29 | 454   | 19.1 | 356  | 5  | Q95P98 |
| 30 | 448   | 18.8 | 436  | 4  | Q8WUR3 |
| 31 | 448   | 18.8 | 436  | 11 | Q8VCM7 |
| 32 | 447.5 | 18.8 | 457  | 11 | Q8R0Z6 |
| 33 | 439.5 | 18.5 | 470  | 4  | Q9BZ20 |
| 34 | 437   | 18.4 | 337  | 6  | Q9BE00 |
| 35 | 435.5 | 18.3 | 235  | 6  | Q28763 |
| 36 | 434.5 | 18.3 | 326  | 6  | Q29042 |
| 37 | 433.5 | 18.2 | 324  | 5  | Q95PA0 |
| 38 | 432   | 18.2 | 279  | 11 | Q9CU50 |
| 39 | 428.5 | 18.0 | 323  | 6  | Q29041 |
| 40 | 427   | 18.0 | 316  | 6  | Q28529 |
| 41 | 423   | 17.8 | 1353 | 13 | Q00546 |
| 42 | 421   | 17.7 | 712  | 4  | Q00531 |
| 43 | 421   | 17.7 | 1358 | 4  | Q92752 |
| 44 | 421   | 17.7 | 1358 | 4  | Q15568 |
| 45 | 420.5 | 17.7 | 307  | 6  | Q28703 |

## ALIGNMENTS

## RESULT 1

Q9EPT7 ID Q9EPT7 PRELIMINARY; PRT; 357 AA.  
AC Q9EPT7;  
DT 01-MAR-2001 (TREMREL. 16, Created)  
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)  
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)  
DE Prothrombinase FGL2  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RA Rycklik D.F., Chien E., Phillippe M.;  
RT "FGL2 Expression in the Sprague-Dawley Rat.";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DDJ databases.  
DR EMBL; AF323608; AAG42269.1; -;  
DR HSSP; P02671; IFRD.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR Pfam; PF00147; fibrinogen\_C; 1.  
DR SMART; SM00186; FBG; 1.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
SQ SEQUENCE 357 AA; 40966 MW; 31183DD9A02EBBA9 CRC64;

Query Match 69.3%; Score 1647; DB 11; Length 357;  
Best Local Similarity 82.7%; Pred. No. 3.9e-100;  
Matches 302; Conservative 24; Mismatches 31; Indels 8; Gaps 1;

QY 75 IEEVFEQVNLKKEEINVLHGRLKLNLMNMIENYVDSKVNLTFFVNSLDGKSKC 134  
Db 1 MEEVFEQVNLKKEEINVLHGRLKLNLMNMIENYVDSKVNLTFFVNSLDGKSKC 52  
QY 135 SEVNLKSELKNAKEEINVLHGRLKLNLMNMIENYVDSKVNLTFFVNSLDGKSKC 194  
Db 53 SOVNLKSELKNAKEEINVLHGRLKLNLMNMIENYVDSKVNLTFFVNSLDGKSKC 112  
QY 195 PSQEQIQSRPQVHLIYKDCSYAIGKRSSSTRYTPDPKNSSEFVYCDMETMGGWTVL 254

```
DB 113 PSQEHQNPVQHLIYKDCSDYYVLGKRSSCIYRVTPDHRNSFEVYCDMETGGWVL 172
QY 255 QARLDGSTNTRTWQDYKAGFNLRRFELGNDKIHLLTKSKEMILRIDLEDFNGVELYA 314
DB 173 QARLDGSTNTRGWQDYKAGFNLRRFELGNDKIHLLTKSKEMILRIDLEDFNGLTLYA 232
QY 315 LYDOFYVANEFKYLRLHVGNTAGDALRFNKHYNHDLKFFTPDKDNDRYPSCNGLY 374
DB 233 YDQFYVANEFKYLRLHVGNTAGDALRFNKHYNHDLKFFTPDKDNDRYPSCNGLY 292
QY 375 YSSGWDFDACLNSANLNGYHYHOKYRGVRNGIFWGTWPGVSAHPGGYKSFKEAKMIRP 434
DB 293 YSSGWDFDACLNSANLNGYHYHOKYRGVRNGIFWGTWPGVSAHPGGYKSFKEAKMIRP 352
QY 435 KHFKP 439
DB 353 KSEKP 357

RESULT 2
Q8WWE4
ID Q8WWE4 PRELIMINARY; PRT; 148 AA.
AC Q8WWE4;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Similar to fibrinogen-like 2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Buthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017813; AAH17813.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 148 AA; 16561 MW; 456B7DBA37C07F21 CRC64;

Query Match 31.5%; Score 750; DB 4; Length 148;
Best Local Similarity 99.3%; Pred. No. 6.4e-42;
Matches 144; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKLANWYLVSSAVLATYGLVAVANNETEELKDERAKDVCVRLSRKCEAGECPYQVS 60
DB 1 MKLANWYLVSSAVLAAYGLVAVANNETEELKDERAKDVCVRLSRKCEAGECPYQVS 60
QY 61 LPPLTIQLPKFSRIEEVFEVONLKEIVNSLKKSCQCKLOADDNDGPRNCLLLPSTG 120
DB 61 LPPLTIQLPKFSRIEEVFEVONLKEIVNSLKKSCQCKLOADDNDGPRNCLLLPSTG 120
QY 121 APGEVGNRVRESEVKNLSSELK 145
DB 121 APGEVGNRVRESEVKNLSSELK 145

RESULT 3
Q9NDQ1
ID Q9NDQ1 PRELIMINARY; PRT; 652 AA.
AC Q9NDQ1;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Fibrinogen-like protein.
GN CI-FIBRN.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,
RA Satou N.;
RT "Characterization of Brachyury downstream notochord genes in the Ciona
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036849; BAB00626.1; -.
DR HSSP; P02671; 1FZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 652 AA; 73252 MW; A492BA325162F0E0 CRC64;

Query Match 26.0%; Score 618.5; DB 5; Length 652;
Best Local Similarity 43.0%; Pred. No. 1.8e-32;
Matches 120; Conservative 53; Mismatches 91; Indels 15; Gaps 6;

QY 171 NYVDSKVANLT-----FVNSLDG-----KCKPCSQEQIQSRPVQHLIYKDCSDYYAIGK 221
DB 371 NYIYSEGSGVINRDGYATNETEGIEFTSRILRATNPTPVQDGGTESLPY-DCAEIYARGV 429
QY 222 RSSEYRVTPDPKNSSEFEVYCDMETGGWTVLQARLDGSTNTRTWQDYKAGFNLRRRE 281
DB 430 RQSGVYDIRPGTK-VTWTVYCDMDTDGGGWTMLQRRIDGIVNFSKGMKSKYKNGEGDINAD 488
QY 282 FWLGNKIHLLT---KSKEMILRIDLEDFNGVELYALYDOFYVANEFKYLRLHVGNYNGT 338
DB 489 HWIGLEKMHIIHSTSNKSRMELINLTDDVSHYANGYVFRISSEKKNYQLIAKRYTGT 548
QY 339 AGDALRFNKHVHNDLKFFTPDKDNDRYPSCNGLYSSGWDFDACLNSANLNGYHYHOKY 398
DB 549 AGDALNVGENYHNLQPTTFDRDNDGYALGNGRYRSGMWFENACFAANLNGNYTGPY 608
QY 399 RGVNGLFWGTWPGVSAHPGGYKSFKEAKMIRPKHF 437
DB 609 QVQNGIYGTWTKLSDS-TSNRYSFKYVDMKVRPLNF 646

RESULT 4
Q8T8A2
ID Q8T8A2 PRELIMINARY; PRT; 220 AA.
AC Q8T8A2;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Fibrinogen-like protein (Fragment).
GN CS-FIBRINOGEN-LIKE.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21920613; PubMed-11923208;
RA Imai K.S., Satou N., Satou Y.;
RT "Early embryonic expression of FGF4/6/9 gene and its role in the
RT Induction of mesenchyme and notochord in Ciona savignyi embryos.";
RL Development 129:1729-1738(2002).
DR EMBL; AB073373; BAB88674.1; -.
FT NON_TER
SQ SEQUENCE 220 AA; 25758 MW; 4160FAE727F3ED06 CRC64;

Query Match 25.0%; Score 593.5; DB 5; Length 220;
Best Local Similarity 49.8%; Pred. No. 1.9e-31;
Matches 107; Conservative 42; Mismatches 61; Indels 5; Gaps 3;

QY 227 YRVTPDPKNSSEFEVYCDMETGGWTVLQARLDGSTNTRTWQDYKAGFNLRRFVLGN 286
DB 4 YELKPN-TDETWMVYCDMETAGGWTVIQKRVGDEENFSRMKRAYKNGFGDKNKHIGL 62
QY 287 DKIHLLT---KSKEMILRIDLEDFNGVELYALYDOFYVANEFKYLRLHVGNYNGTAGDAL 343
RN [1]
```

```

Db 63-ERMHLTTSNKRRLKRLIDLWDVNVHAYETPRVRGEGKNYQLIAKFKSTAGDAL 122
QY 344 RENKYNHDLAFFTPDKNDYRPSGNGCLYYSGWDFDACL SANLNGKYYHOKYRGVRN 403
Db 123 NGENYNHMAQFTFDNDGALNGCRYYRSGWDFNACFAANLNGYTGPRGVON 182
QY 404 GFWGTWPGVSEAHPPGYKSSFKKAKMIRPKHF 438
Db 183 GIWGTWYKLSDRSNA-RYSEKYVDMKVRPLNFE 216

RESULT 5
Q9D2D2
ID Q9D2D2 PRELIMINARY; PRT; 496 AA.
AC Q9D2D2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 11 days pregnant adult female ovary and uterus cDNA, RIKEN full-length
DE enriched library, clone:5031400E18, full insert sequence (Angiopoietin
DE 2).
AGP2.
Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=OVARY, AND UTERUS;
RX MEDLINE=21083660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki K., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RA Nature 409:685-690(2001).
[2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK019860; BAB31887.1; -.
DR EMBL; BC027216; AAH27216.1; -.
DR HSSP; P02671; 1FZD.
DR MGD; MGI:1202890; Agpt2.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 496 AA; 56575 MW; E7563B498A0EF31 CRC64;

Query Match 22.1%; Score 526; DB 11; Length 496;
Best Local Similarity 38.9%; Pred. No. 1.4e-26;
Matches 127; Conservative 66; Mismatches 153; Indels 40; Gaps 9;

QY 61 LPLTITQLPKQFSEIEFKEVONLKEIVASLKSCQDCKLOADNGDPCGNCLLLPTG 120
Db 142 LTDVEAQLVQNTRL-----ELQLQHSISTNKLEKQLDITSEINKLQNKNSFL----- 191
QY 121 APGEVDGNRVRELESVKNLSSELKNAKEEINVHLGR-----EKLNLVNMNNIENY 172

```

```

Db 192 -----EOKVLDEGKHSQLOSMEKQDEQLVLSKQSSVIDELEKLYTATVNN--SL 243
QY 173 VDSKVANLTFVWNSLDGKSCPKSQEQIQSRPVQVHLYIKDCSDYYAIGKRSSEYRVTPD 232
Db 244 LQKQCHDLMTVNSLLTMMSSPNKSSVAIRKEEQTFPRDCAEIFKSLGTLTSGIYTLTFP 303
QY 233 PKNSSFEVYCDMETMGWTVLQARLDGSTNFTRWQDYKAGFGNLRREFWLGNDKIHLL 292
Db 304 NSTEIRAICDMVGWGVTVJQHRDGSVDFTWKEYKGFSGPLGEYWLGNFEVSOL 363
QY 293 TKSSEMILRLDLEDFNGVELYALYDQFYVANEFLKYLRLHVGNYNGTAGDALRFNKHYNH 352
Db 364 TCQHRVYLKQLKDWECNEAHSYDFYLAGESNRYIHLTGLTGTAGKLSSISQPGSD- 422
QY 353 LKFFTPDKNDYRPSGNGCLYYSGWDFDACL SANLNGKYYHOKYRGVR-NGIFWGTWP 411
Db 423 ---FSTKDSNDKIC- KCSQMLSGWDFDAGPSNLNGYYPOKQNTNFKENGIRKYYWK 478
QY 412 GYSEAHPPGYKSSFKKAKMIRPKHF 437
Db 479 G-----SGY--SLKATTWMIKPADF 496

RESULT 6
Q9DER1
ID Q9DER1 PRELIMINARY; PRT; 407 AA.
AC Q9DER1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Angiopoietin-2B.
GN ANGIOPOIETIN-2B.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422311; PubMed=10964717;
RA Mezquita J., Mezquita P., Montserrat P., Mezquita B., Francome V.,
RA Villagrasa X., Mezquita C.;
RT "Genomic structure and alternative splicing of chicken angiopoietin-
RT 2.";
RL Biochem. Biophys. Res. Commun. 275:643-651(2000).
DR EMBL; AJ289778; CAC08175.1; -.
DR HSSP; P02671; 1FZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 407 AA; 46687 MW; 8863BA0A8C7A41C CRC64;

Query Match 22.0%; Score 524; DB 13; Length 407;
Best Local Similarity 38.3%; Pred. No. 1.5e-26;
Matches 125; Conservative 45; Mismatches 120; Indels 36; Gaps 10;

QY 127 DNRVRELESVKNLSSELKNAKEEINVHLGR-----LEK-LNLVNMNNI-----ENVY 173
Db 103 EKRYLEMDKHTLQLKSIRKDEKQLVLRONSIIIELEKQLVATVNSVLQKQCHDL 162
QY 174 DSKVAN-LTFVWNSLDGKSCPKSQEQIQSRPVQVHLYIKDCSDYYAIGKRSSEYRVTPD 232
Db 163 METVHLLTMTSPNSAKKNFLAKEEQIS-----FKDCAEAFKSLTSGIYTLTFP 214
QY 233 PKNSSFEVYCDMETMGWTVLQARLDGSTNFTRWQDYKAGFGNLRREFWLGNDKIHLL 292
Db 215 NSAQEKAYKAYCDMESNGWGVTVLQRRDGSVDFTWKEYKGFDPAGEYWLGNFEVSOL 274
QY 293 TKSSEMILRLDLEDFNGVELYALYDQFYVANEFLKYLRLHVGNYNGTAGDALRFNKHYNH 352
Db 275 TNQRYVYLKILKDWEGNEATLYDQFYLANEQRYIHLKGLTGTAGKLSSISQPGND- 333

```

```

QY 353 LKFTTTPDKNDNRYPSGNCGLIYSSGWFFDACLSANLNGKYYHOKYRGVR-NGIFWGTWP 411
Db 334 ---FSTKADADNDKIC-KCSOMLTGGWFFDAGPSNLNGMYPLRQNNKFKNGIKWYWK 389
QY 412 GVSEAHPGGKYSFKEAKMIRPKHF 437
Db 390 G-----SGY--SLKATTMIRPADF 407

RESULT 7
Q9DER0
ID Q9DER0 PRELIMINARY; PRT: 441 AA.
AC Q9DER0;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Angiopoietin-2.
GN Angiopoietin-2C.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422311; PubMed=10964717;
RA Mezquita J., Mezquita P., Montserrat P., Mezquita B., Francone V.,
RA Villagrasa X., Mezquita C.;
RT "Genomic structure and alternative splicing of chicken angliopoietin-
RT 2.";
RL Biochem. Biophys. Res. Commun. 275:643-651(2000).
DR EMBL; AJ289777; CAC08174.1; -.
DR HSP; P02671; IYZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 441 AA; 50472 MW; DC98127FE0CE34B2 CRC64;

Query Match 22.0%; Score 524; DB 13; Length 441;
Best Local Similarity 38.3%; Pred. No. 1.6e-26;
Matches 125; Conservative 45; Mismatches 120; Indels 36; Gaps 10;

QY 127 DNRVRESEVKNLSSELKNAKEEINVLR-----LEK-LNLVNMNI-----ENVV 173
Db 137 EKRYLEMDKHTLQLSIKDEKDLQVLVARQNSIIELEKQLVTATVNSVLOKQOHD 196
QY 174 DSKVAN-LTFVNSLDGKCKPQSOIQSRPVQHLIYKDCSDYVYAIKRSSETRYVTPD 232
Db 197 METHVNLITWISTPNSAKNFIKAEQIS-----FKDCAEAFKSGLTTSGLTYLTFF 248
QY 233 PKNSSEVYCDMETGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRFELGNDKIHLL 292
Db 249 NSAQEKKAYCDMESGGGWTVLQRRDGSVDHFRTWKEYKIGFGDPAGEYWLGNFVSQ 308
QY 293 TKSEMTLRDLDEDFNGVELYALDYFVANEFLKYLRLHVGNYNGTAGDALRNFKNHYND 352
Db 309 TNOKRYVLKILKDWEGNEAYTLDYQFLANEQYRIHLKGLTGTAGKISSISQPGND- 367
QY 353 LKFTTTPDKNDNRYPSGNCGLIYSSGWFFDACLSANLNGKYYHOKYRGVR-NGIFWGTWP 411
Db 368 ---FSTKADADNDKIC-KCSOMLTGGWFFDAGPSNLNGMYPLRQNNKFKNGIKWYWK 423
QY 412 GVSEAHPGGKYSFKEAKMIRPKHF 437
Db 424 G-----SGY--SLKATTMIRPADF 441

RESULT 8
Q9DER2
ID Q9DER2 PRELIMINARY; PRT: 493 AA.
AC Q9DER2;

```

```

DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Angiopoietin-2.
GN Angiopoietin-2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422311; PubMed=10964717;
RA Mezquita J., Mezquita P., Montserrat P., Mezquita B., Francone V.,
RA Villagrasa X., Mezquita C.;
RT "Genomic structure and alternative splicing of chicken angliopoietin-
RT 2.";
RL Biochem. Biophys. Res. Commun. 275:643-651(2000).
DR EMBL; AJ289777; CAC08174.1; -.
DR HSP; P02671; IYZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 493 AA; 56393 MW; BDIA21F90172F6DA CRC64;

Query Match 22.0%; Score 524; DB 13; Length 493;
Best Local Similarity 38.3%; Pred. No. 1.9e-26;
Matches 125; Conservative 45; Mismatches 120; Indels 36; Gaps 10;

QY 127 DNRVRESEVKNLSSELKNAKEEINVLR-----LEK-LNLVNMNI-----ENVV 173
Db 189 EKRYLEMDKHTLQLSIKDEKDLQVLVARQNSIIELEKQLVTATVNSVLOKQOHD 248
QY 174 DSKVAN-LTFVNSLDGKCKPQSOIQSRPVQHLIYKDCSDYVYAIKRSSETRYVTPD 232
Db 249 METHVNLITWISTPNSAKNFIKAEQIS-----FKDCAEAFKSGLTTSGLTYLTFF 300
QY 233 PKNSSEVYCDMETGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRFELGNDKIHLL 292
Db 301 NSAQEKKAYCDMESGGGWTVLQRRDGSVDHFRTWKEYKIGFGDPAGEYWLGNFVSQ 360
QY 293 TKSEMTLRDLDEDFNGVELYALDYFVANEFLKYLRLHVGNYNGTAGDALRNFKNHYND 352
Db 361 TNOKRYVLKILKDWEGNEAYTLDYQFLANEQYRIHLKGLTGTAGKISSISQPGND- 419
QY 353 LKFTTTPDKNDNRYPSGNCGLIYSSGWFFDACLSANLNGKYYHOKYRGVR-NGIFWGTWP 411
Db 420 ---FSTKADADNDKIC-KCSOMLTGGWFFDAGPSNLNGMYPLRQNNKFKNGIKWYWK 475
QY 412 GVSEAHPGGKYSFKEAKMIRPKHF 437
Db 476 G-----SGY--SLKATTMIRPADF 493

RESULT 9
Q9P2Y7
ID Q9P2Y7 PRELIMINARY; PRT: 495 AA.
AC Q9P2Y7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Angiopoietin-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99126459; PubMed=9927494;
RA Tanaka S., Mori M., Sakamoto Y., Makuuchi M., Sugimachi K.,
RA Wands J.R.;
RT "Biologic significance of angliopoietin-2 expression in human

```





```
Db 197 EQCLRFSDTHVSPLVQVWPHIPNSQQYTGLLGGNGEIORDPGYPRDLMPDPLA 256
QY 209 -----IYKCDSDYYAIGKRSETYRTVPDPKNSSEFYCDMETNGG 249
Db 257 TSPTKSPFKTPPVTFETNECPKQQAQKAGHSGVGIYIKPENSQPMQLWCNSLDPG 316
QY 250 GWTVLOARLDGSTNFTRTWODYKAGFCNLRREFWLGNDKIHLLTKSKEMILRLDLEDFNG 309
Db 317 GWTVIQRTDGSVNFERNWENYKKGFGNDGEYWLGLENIYMSNQDNKLLILELWSD 376
QY 310 VELYALYDQFYVANEFLYRLHVGNCTAGDALRNFKNHYHDLKFTFTDPDKNDRYPSSG 369
Db 377 KKVTAESYSPLEPESEFYRLRLGTQYGNAGDSMMW-----HNGKQFTTLDKDMY-AG 430
QY 370 NCGLYSYSGWDFDACLNSANLNGKYYH-QKYRGV-RNGIFWCTWPGVSEAHPGGYKSSPKE 427
Db 431 NCAHFHKGWYNACAHNSLNCVYRGCHYRSKHQDGIFFWAEYRG-----GSY--SLRA 482
QY 428 AKMIRP 434
Db 483 VQMIKP 489
```

## RESULT 14

```
Q9HBP3
ID Q9HBP3 PRELIMINARY; PRT; 244 AA.
AC Q9HBP3:
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 27.8 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel Human cDNA clones with function of inhibiting cancer cell
growth";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF218015; AAC17257.1; -.
DR HSSP; P02671; IYZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
SQ SEQUENCE 244 AA; 27825 MW; E9D0ACBF87A7139E CRC64;
```

Query Match 20.9%; Score 496; DB 4; Length 244;  
Best Local Similarity 38.9%; Pred. No. 5.1e-25;  
Matches 102; Conservative 43; Mismatches 95; Indels 22; Gaps 6;

```
QY 177 VANLFVYNSLDGKCKSPQSQIQSRPVQHLYIKDCSDYYAIGKRSETYRTVPDPKNS 236
Db 4 VNNLLTMSTNSKDPVAKREQIS-----FRDCAEVFKSGHTINGIYTIPTFNSTE 55
QY 237 SEFYCDMETNGGWTVLOARLDGSTNFTRTWODYKAGFCNLRREFWLGNDKIHLLTKSK 296
Db 56 EIKAYCDMEAGGGGWTIIORREDGSDVQRTWKKEYKVGFGPSGNGYFVSLNQNOQ 115
QY 297 EMILRLDLEDFNGVELYALYDQFYVANEFLYRLHVGNCTAGDALRNFKNHYHDLKFP 356
Db 116 RYVLKHLKDWEGNEAYSLEYEHFLLSSEELNYRLHLKGLTGTAGKISSIQPCND---F 171
QY 357 TTPDKNDRYPSSGNGCYLYSSGWDFDACLNSANLNGKYYHQRVGRV-NGIFWCTWPGVSE 415
Db 172 STKGDNDKIC-KCSQMLTGGWDFDAGCPGSLNNGMYIPQRTNKNFKNGIKWYKWG--- 227
```

```
QY 416 AHGGYKSSKKEAKMMIRPKHF 437
Db 228 ---SGY--SLKATTMIRPADF 244

RESULT 15
Q90Z19
ID Q90Z19 PRELIMINARY; PRT; 513 AA.
AC Q90Z19:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Angiopoietin-1.
GN ANGI.
OS Brachydanio rerio (Zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21391693; PubMed=11500985;
RA Pham V.N., Roman B.L., Weinstein B.M.;
RT "Isolation and expression analysis of three zebrafish angiopoietin
RT genes.";
RL Dev. Dyn. 221:470-474(2001).
DR EMBL; AF379602; AAK83347.1; -.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR PRINTS; PR01608; BACINVASINC.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; UNKNOWN.1.
SQ SEQUENCE 513 AA; 58360 MW; 0627777A39847DBB CRC64;
```

Query Match 20.7%; Score 491.5; DB 13; Length 513;  
Best Local Similarity 26.6%; Pred. No. 2.6e-24;  
Matches 142; Conservative 86; Mismatches 179; Indels 127; Gaps 16;

```
QY 8 WLSSAVIATYGFVLVANNETETEEKDERAKDVCPRLESRGKCE---PAGECPYQVSLPPL 64
Db 2 WMGCLFLTA--ALLVWADCGVEQKTLGMSSTPKSSSGRRYHRIHQHOGCSTFILPES 59
QY 65 TQLPKQF-----SRIEEVFKE-VQNLEKIVNSLKKSC 96
Db 60 DCMTCREFKSGTAYANALORDAPQPEADLSNKKIQOLEHVWENYTWLQKTIENYIKNM 119
QY 97 QDCKLQADDNGDPCRNGLLLPSTCAPGEVGDN-----RVRELESEV----- 137
Db 120 KTEMVQLQSQSAVHNHTAAML-----EMGTSLLSOTAEQTRKLTDTVETVLNQTSLREI 172
QY 138 -----NKLSELNKAKEINVLH---GRLE-----KLNLMNMNNIENYVDS 175
Db 173 QLLENSLSTNKLKQLMIQINEINKIHDKNGFLEKKMOELEDRHQELESRLRTEKSDLQA 232
QY 176 KYANLTFVNSLDGKCKSPQSQIQSRPVQHLYI----- 209
Db 233 LVSRQSSVIRELENQSLRATGNSTALQRQDDLMESMRSLSLCAKDAATAVEPNTKQA 292
QY 210 -----YKDCSDYYAIGKRSETYRTVPDPKNSSEFYCDMETNGGWTVLOARLDGSTN 264
Db 293 DEERKFRDADLYOAGFQKNGVYITINISPOETK-KVYCVMSAGSGGWTVIQKREDGTVDF 351
QY 265 TPTWQDYKAGFCNLRREFWLGNDKIHLLTKSKEMILRLDLEDFNGVELYALYDQFYVANE 324
Db 352 QKTWKKEYKMGFGSVSGEHWLGNFVHVLTNRQHLRVLSDWGDHQAQFQSYDSFHIDSE 411
QY 325 FLKYRLHVGNCTAGDALRNFKNHYHDLKFTFTDPDKNDRYPSSGNGCYLYSSGWDFDACL 384
Db 412 KOKYRLFLKTHSGTAGRQSSLAHV-GAD---FSTKQVDNDNC-TCKCALMLSGGWWYDAC 466
QY 385 LSNANLNGKYYH-QYRGVNGIFWCTWPGVSEAHPGGYKSSKKEAKMMIRPKHF 437
Db 467 GPSNLNGVYRQGHVKGKFNKIKWHYFKGFS-----YSLSRTVMIRGADF 512
```

Mon Dec 2 07:10:24 2002

us-09-902-563-2.rspt

Page 8

Search completed: December 2, 2002, 06:53:23  
Job time : 122.903 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:38:29 ; Search time 61.8855 Seconds  
(without alignments)  
945.247 Million cell updates/sec

Title: US-09-902-563-2

Perfect score: 2378

Sequence: 1 MLLANWYLLSSAVLATYGFLL.....GYKSSPKAKMIRKHKPK 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description                  |
|------------|--------|-------------|--------|-------|------------------------------|
| 1          | 2378   | 100.0       | 439    | 20    | AAW88235 Human prothrombina  |
| 2          | 1853.5 | 77.9        | 432    | 20    | AAW88236 Mouse prothrombina  |
| 3          | 524    | 22.0        | 314    | 15    | AAW65759 Rat hepatic parench |
| 4          | 522    | 22.0        | 496    | 17    | AAW01411 Human TIE-2 ligand  |
| 5          | 522    | 22.0        | 496    | 17    | AAW94605 Human TIE-2 ligand  |
| 6          | 522    | 22.0        | 496    | 19    | AAW47532 Human TIE-2 ligand  |
| 7          | 522    | 22.0        | 496    | 21    | AAW82839 Human angioipoietin |
| 8          | 522    | 22.0        | 496    | 21    | AAW78903 Human angioipoietin |
| 9          | 522    | 22.0        | 496    | 21    | AAW78906 Human angioipoietin |
| 10         | 522    | 22.0        | 496    | 23    | AAW77944 Amino acid sequenc  |

|    |       |      |      |    |          |                     |
|----|-------|------|------|----|----------|---------------------|
| 11 | 522   | 22.0 | 572  | 22 | ABG06771 | Novel human diagno  |
| 12 | 522   | 22.0 | 1033 | 22 | ABG11614 | Novel human diagno  |
| 13 | 521.5 | 21.9 | 498  | 19 | AAW47528 | Amino acid sequenc  |
| 14 | 520   | 21.9 | 444  | 23 | AAW77252 | Human angioipoietin |
| 15 | 517.5 | 21.8 | 499  | 19 | AAW47529 | Amino acid sequenc  |
| 16 | 506.5 | 21.3 | 491  | 21 | AAW48000 | Hepatocyte prolife  |
| 17 | 506   | 21.3 | 312  | 17 | AAW94317 | Hepatocyte prolife  |
| 18 | 506   | 21.3 | 339  | 21 | AAW43637 | Human cancer assoc  |
| 19 | 505.5 | 21.3 | 312  | 17 | AAW94316 | Hepatocyte prolife  |
| 20 | 505   | 21.2 | 491  | 20 | AAW26196 | Human zapo3 protei  |
| 21 | 505   | 21.2 | 491  | 20 | AAW05397 | Human TIE ligand N  |
| 22 | 505   | 21.2 | 491  | 21 | AAW47996 | Angioipoietin-like  |
| 23 | 505   | 21.2 | 491  | 21 | AAW24393 | Human PRO188 prote  |
| 24 | 505   | 21.2 | 491  | 21 | AAW69483 | Amino acid sequenc  |
| 25 | 505   | 21.2 | 491  | 21 | AAW68761 | An angioipoietin-re |
| 26 | 505   | 21.2 | 491  | 22 | AAU12310 | Human PRO188 polyp  |
| 27 | 505   | 21.2 | 491  | 22 | AAW60340 | Human neovasculari  |
| 28 | 505   | 21.2 | 491  | 22 | AAW53067 | Human angiogenesis  |
| 29 | 505   | 21.2 | 491  | 23 | AAW19826 | Human TIE ligand N  |
| 30 | 504.5 | 21.2 | 312  | 15 | AAW65760 | Human hepatic pare  |
| 31 | 500   | 21.0 | 491  | 21 | AAW44841 | Human secreted pro  |
| 32 | 490   | 20.6 | 689  | 21 | AAW96780 | Ang-2-FD-Fc-FD fus  |
| 33 | 489   | 20.6 | 346  | 20 | AAW23736 | Human AR-1 protein  |
| 34 | 489   | 20.6 | 346  | 20 | AAW05399 | Human TIE ligand N  |
| 35 | 489   | 20.6 | 346  | 21 | AAW90400 | Human angioipoietin |
| 36 | 489   | 20.6 | 346  | 21 | AAW33432 | Human PRO356 prote  |
| 37 | 489   | 20.6 | 346  | 21 | AAW24408 | Human PRO356 prote  |
| 38 | 489   | 20.6 | 346  | 21 | AAW95347 | Human PRO356 antit  |
| 39 | 489   | 20.6 | 346  | 22 | AAW79032 | Human protein SEQ   |
| 40 | 489   | 20.6 | 346  | 22 | AAW73468 | Human CDT6 (cornea  |
| 41 | 489   | 20.6 | 346  | 22 | AAW20116 | Human immunostimul  |
| 42 | 489   | 20.6 | 346  | 22 | AAW50911 | Human PRO356 prote  |
| 43 | 489   | 20.6 | 346  | 22 | AAW50983 | Human PRO356 prote  |
| 44 | 489   | 20.6 | 346  | 22 | AAW53068 | Human angiogenesis  |
| 45 | 489   | 20.6 | 686  | 21 | AAW96778 | Ang-2-FD-FD-Fus     |

ALIGNMENTS

|          |  |                                     |
|----------|--|-------------------------------------|
| RESULT 1 | AAW88235   | AAW88235 standard; Protein; 439 AA. |
| ID       | AAW88235;  |                                     |
| XX       | AAW88235;  |                                     |
| AC       | AAW88235;  |                                     |
| DT       | 15-MAR-1999 (first entry)  |                                     |
| XX       | Human prothrombinase Fgl2 protein.                                 |                                     |
| DE       | Prothrombinase; hfgl2; Fgl2; human; immune coagulation; antibody;  |                                     |
| KW       | inhibitor; infection; graft rejection; glomerulonephritis; cancer; |                                     |
| KW       | gastrointestinal disease; foetal loss; therapy; vaccine.           |                                     |
| XX       | Homo sapiens.  |                                     |
| OS       | Homo sapiens.  |                                     |
| PH       | Key  | Location/Qualifiers                 |
| FT       | Modified-site  | 179...183                           |
| FT       | Modified-site  | /label= Asn is N-glycosylated       |
| FT       | Modified-site  | 235...238                           |
| FT       | Modified-site  | /note= "Asn is N-glycosylated"      |
| FT       | Modified-site  | 262...265                           |
| FT       | Modified-site  | /note= "Asn is N-glycosylated"      |
| FT       | Modified-site  | 336...337                           |
| FT       | Modified-site  | /note= "Asn is N-glycosylated"      |
| FT       | Domain   | 213...439                           |
| FT       | Peptide  | /note= "fibrinogen related domain"  |
| FT       | Peptide  | 364...378                           |
| FT       | Peptide  | /note= "epitope (Claim 4)"          |
| XX       | W09851335-A1.  |                                     |
| PN       | W09851335-A1.  |                                     |
| XX       | 19-NOV-1998.   |                                     |
| PD       | 19-NOV-1998.   |                                     |

```

XX PF 15-MAY-1998; 98WO-CA00475.
XX PR 10-OCT-1997; 97US-0061684.
XX PR 15-MAY-1997; 97US-0046537.
XX PA (LEVY/) LEVY G.
XX PI Levy G;
XX DR WPI; 1999-059687/05.
XX DR N-PSDB; AAV84139.
XX PT Modulating immune coagulation - by using Fgl2 antibodies and
XX PT compounds, used to treat conditions including graft rejection and
XX PT foetal loss
XX PS Claim 8; Page 66-67; 105pp; English.
XX PR This is the amino acid sequence of human prothrombinase Fgl2, as
XX PR predicted from hfgl2 DNA (see AAV84139). Fgl2 is a 70 kDa
XX PR transmembrane serine protease that has immune procoagulant activity.
XX PR The invention provides a method for inhibiting immune coagulation by
XX PR inhibiting the activity or expression of Fgl2. The method can be
XX PR used in vivo to treat a condition which requires a reduction in
XX PR immune coagulation such as bacterial and viral infections, cancer,
XX PR glomerulonephritis, a number of gastrointestinal diseases,
XX PR allograft and xenograft rejection and foetal loss. An Fgl2-specific
XX PR antibody, an Fgl2 antisense oligonucleotide, or a substance that
XX PR affects prothrombinase activity of a Fgl2 protein may be used to
XX PR treat a condition requiring a reduction in procoagulant activity.
XX PR A vaccine containing an Fgl2 protein or peptide is used for
XX PR prevention of graft rejection or foetal loss (claimed).
XX SQ Sequence 439 AA;

Query Match 100.0%; Score 2378; DB 20; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.9e-205;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLANWYLSAVLATYGLVANNTEETIKDERAKDVCVRLESRGKCEAGECPYQVS 60
DB 1 MKLANWYLSAVLATYGLVANNTEETIKDERAKDVCVRLESRGKCEAGECPYQVS 60
QY 61 LPPLTQLPKQFSRIEEVFKEVQNLKEIYVNSLKKSCQDCKLQADDNGDGRNGLLLPSTG 120
DB 61 LPPLTQLPKQFSRIEEVFKEVQNLKEIYVNSLKKSCQDCKLQADDNGDGRNGLLLPSTG 120
QY 121 APGEVGDNRVRELESEVNKLSSELKNAKEEINVLHGRLEKLNLYNMNNTIENVDSKVANL 180
DB 121 APGEVGDNRVRELESEVNKLSSELKNAKEEINVLHGRLEKLNLYNMNNTIENVDSKVANL 180
QY 181 TFVNSLDGKCKSPQSQISRPVQHLKYKDCSDYVAGKRSSEYRYTPDPKNSFEV 240
DB 181 TFVNSLDGKCKSPQSQISRPVQHLKYKDCSDYVAGKRSSEYRYTPDPKNSFEV 240
QY 241 YCDMETMGSGWTVLQARLDGSTNFTRTWODYKAGFNLRRFVLGNDKTHLLTKSEKIL 300
DB 241 YCDMETMGSGWTVLQARLDGSTNFTRTWODYKAGFNLRRFVLGNDKTHLLTKSEKIL 300
QY 301 RIDLEDFNGVELYALDYVANEFLKYRLHVGNVNGTAGDALRFNKHVNHDLKFFTPD 360
DB 301 RIDLEDFNGVELYALDYVANEFLKYRLHVGNVNGTAGDALRFNKHVNHDLKFFTPD 360
QY 361 KNDRTYPSNGCLYYSGSGWFDACLSANLNGKYHQQYRGVNRGIFWGTWFGVSEAHPGG 420
DB 361 KNDRTYPSNGCLYYSGSGWFDACLSANLNGKYHQQYRGVNRGIFWGTWFGVSEAHPGG 420
QY 421 YKSFKEAKWMLRPKHEKP 439
DB 421 YKSFKEAKWMLRPKHEKP 439

```

```

RESULT 2
AAW88236
ID AAW88236 standard; Protein; 432 AA.
XX AC AAW88236;
XX DT 15-MAR-1999 (first entry)
XX DE Mouse prothrombinase Fgl2 protein.
XX KW Prothrombinase; Fgl2; mouse; immune coagulation; antibody;
XX KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;
XX KW gastrointestinal disease; foetal loss; therapy; vaccine.
XX OS Mus sp.
XX FT Key Location/Qualifiers
XX FT Modified-site 172..174
XX FT /label= Asn is N-glycosylated
XX FT Modified-site 228..231
XX FT /note= "Asn is N-glycosylated"
XX FT Modified-site 256..259
XX FT /note= "Asn is N-glycosylated"
XX FT Modified-site 323..325
XX FT /note= "Asn is N-glycosylated"
XX FT Domain 213..439
XX FT /note= "fibrinogen related domain"
XX PN W09851335-A1.
XX PR 19-NOV-1998.
XX PR 15-MAY-1998; 98WO-CA00475.
XX PR 10-OCT-1997; 97US-0061684.
XX PR 15-MAY-1997; 97US-0046537.
XX PA (LEVY/) LEVY G.
XX PI Levy G;
XX DR WPI; 1999-059687/05.
XX DR N-PSDB; AAV84140.
XX PT Modulating immune coagulation - by using Fgl2 antibodies and
XX PT compounds, used to treat conditions including graft rejection and
XX PT foetal loss
XX PS Claim 8; Page 70-71; 105pp; English.
XX PR This is the amino acid sequence of mouse prothrombinase Fgl2, as
XX PR predicted from fgl2 DNA (see AAV84140). Fgl2 is a 70 kDa
XX PR transmembrane serine protease that has immune procoagulant activity.
XX PR The human Fgl2 amino acid sequence is given in AAW88236. The
XX PR invention provides a method for inhibiting immune coagulation by
XX PR inhibiting the activity or expression of Fgl2. The method can be
XX PR used in vivo to treat a condition which requires a reduction in
XX PR immune coagulation such as bacterial and viral infections, cancer,
XX PR glomerulonephritis, a number of gastrointestinal diseases,
XX PR allograft and xenograft rejection and foetal loss. An Fgl2-specific
XX PR antibody, an Fgl2 antisense oligonucleotide, or a substance that
XX PR affects prothrombinase activity of a Fgl2 protein may be used to
XX PR treat a condition requiring a reduction in procoagulant activity.
XX PR A vaccine containing an Fgl2 protein or peptide is used for
XX PR prevention of graft rejection or foetal loss (claimed).
XX SQ Sequence 432 AA;

Query Match 77.9%; Score 1853.5; DB 20; Length 432;
Best Local Similarity 77.7%; Pred. No. 3.4e-158;
Matches 341; Conservative 42; Mismatches 49; Indels 7; Gaps 4;
QY 1 MKLANWYLSAVLATYGLVANNTEETIKDERAKDVCVRLESRGKCEAGECPYQVS 60

```

Db 1 MRLPCWLLSSAVLAACR-AVEHNLTEGLSDASQAACPABLESGRC-EGSQCPFLT 58  
 QY 61 LPPLTIQLPKOFSRIEVEKVEQNLKEIVNSLKKSCQDCKLOADDGDPGRNGLLPSTG 120  
 Db 59 LPPLTIQLPKOFSRIEVEKVEQNLKEIVNSLKKSCQDCKLOADDGDPGRNGLLPSTG 114  
 QY 121 APCEVGDNRVRELESEVNKLSELKNAKEEINVLHGRLEKLNLYNMNNIENYVDSKVANL 180  
 Db 115 AE-TAEDSRVQELSEVNKLSELKNAKEEINVLHGRLEKLNLYNMNNIENYVDSKVANL 173  
 QY 181 TFWNSLDGKSCPCSOBOIQSRPVQHLIYKDCSDYYAIGKRSSEYRVTPDPKNSFEV 240  
 Db 174 TFWNSLDGKSCPCSOBOIQSRPVQHLIYKDCSDYYAIGKRSSEYRVTPDPKNSFEV 233  
 QY 241 YCDMETMGWTVLQARLDGSTNFTRTWQDYKAGFCNLRREFWLGNDKIHLLTKSEMIL 300  
 Db 234 YCDMETMGWTVLQARLDGSTNFTRTWQDYKAGFCNLRREFWLGNDKIHLLTKSEMIL 293  
 QY 301 RIDLEDFNGVELYALYDQFYVANEELKYRLHVGNYNGTAGDALRPNKHNDLKFTTTPD 360  
 Db 294 RIDLEDFNGVELYALYDQFYVANEELKYRLHVGNYNGTAGDALRPNKHNDLKFTTTPD 353  
 QY 361 KNDRYPSGNGGLYSSGWMFDCALSANLNGKYYHOKYRGVRNGIFWGTWPGVSEAHFGG 420  
 Db 354 KNDRYPSGNGGLYSSGWMFDCALSANLNGKYYHOKYRGVRNGIFWGTWPGVSEAHFGG 413  
 QY 421 YKSSPEAKMMIRPKHFKP 439  
 Db 414 YKSSPEAKMMIRPKHFKP 432

## RESULT 3

AAR65759  
 ID AAR65759 standard; protein; 314 AA.

XX AAR65759;

XX 17-MAY-1995 (first entry)

XX Rat hepatic parenchymal cell growth factor.

XX Hepatic parenchymal cell growth factor; HPGF; liver diseases;  
 XX liver cancer; cirrhosis.

XX Rattus rattus.

XX Key Location/Qualifiers  
 XX Peptide 1..24  
 XX /label= sig\_peptide

PN WO9421678-A.

XX 29-SEP-1994.

XX 22-MAR-1994; 94WO-JP00455.

XX 23-MAR-1993; 93JP-0063905.

XX (TAIS) TAISHO PHARM CO LTD.

PI Hanada K, Hara H, Matsuki Y, Shindo S, Yoshimura H;

XX WPI; 1994-316940/39.

XX N-PSDB; AAQ77818.

XX Hepatic parenchymal cell growth promoter peptide - is isolated  
 XX from human or animal liver cell or produced by recombinant  
 XX techniques and used for therapy of liver diseases

XX Claim 3; Page 34; 47pp; Japanese.

XX AAQ77818 encodes AAR65759 rat hepatic parenchymal cell growth factor

CC (HPGF). The protein or the N-terminal peptide (AAR65759) may be used in  
 CC the diagnosis and treatment of liver diseases, such as liver cancer and  
 CC cirrhosis.

SQ Sequence 314 AA;

Query Match 22.0%; Score 524; DB 15; Length 314;  
 Best Local Similarity 37.6%; Pred. No. 1.4e-38;  
 Matches 124; Conservative 46; Mismatches 104; Indels 56; Gaps 13;

QY 125 VGD-NRVRE---LESEVNKLSELKNAKEEINVLHGRLEKLNLYNMNNIENYVDSKVAN 179

Db 23 LGDENCLQEVRLEAQRVQLETRVKKQVVIQALLHEK--EVQFLDRGQEDSFID----- 75

QY 180 LTFVNSLDGKSCPCSOBOIQSRPVQHLIYKDCSDYYAIGKRSSEYRVTPDPKNSFE 239

Db 76 -----LQGR-----RH--YADCSEIYNDGFKHSGFYKIKPQSLAEFS 111

QY 240 VYCDMETMGWTVLQARLDGSTNFTRTWQDYKAGFCNLRREFWLGNDKIHLLTKSK 296

Db 112 VYCDMSD--GGGWTVIQRRSDGSENFNRGWNDEYNGFNGVQSGEYWLGNKNINLITWQG 170

QY 297 EMILRIDLEDFNGVELYALYDQFYVANEELKYRLHVGNYNGTAGDALRPNKH-----YN 350

Db 171 DYTLLKIDLTDFEKNRFAQYKFKVGDSEKSYELNIGESYGTAGDSLSGTFHPEVQWNAS 230

QY 351 HDLKFTTTPDKNDRYPSGNGGLYSSGWMFDCALSANLNGKYYHOKYRG--VRNGIFWGT 409

Db 231 HOTMKFSTRDRNDNY--NGNCAEEQSGWMFNCHSANLNGYTYQGPYRAETDNGVWYWT 289

QY 410 WPGVSEAHFGGKSSPEAKMMIRPKHFKP 439

Db 290 W-----RGWYSLKSVVMKIRPSDFIP 311

## RESULT 4

AAW01411  
 ID AAW01411 standard; Protein; 496 AA.

XX AAW01411;

DT 11-FEB-1997 (first entry)

XX Human TIE-2 ligand 2.

XX TIE-2 ligand 2; tyrosine kinase with Ig and EGF homology domain;  
 KW receptor; antagonist; neovascularisation; wound healing; ischaemia;  
 KW leukopenia; thrombocytopenia; anaemia; angiogenesis; tumour;  
 KW atherosclerosis; inflammation; diagnosis; therapy.

OS Homo sapiens.

XX WO9631598-A1.

XX 10-OCT-1996.

XX 05-APR-1996; 96WO-US04806.

XX 06-OCT-1995; 95WO-US12935.

PR 06-APR-1995; 95US-0418595.

XX (REGE-) REGENERON PHARM INC.

XX Aldrich TM, Bruno J, Davis S, Goldfarb M, Jones PF;

PI Maisonnier PC, Radziejewski C, Yancopoulos GD;

XX WPI; 1996-465021/46.

DR N-PSDB; AAT44321.

XX TIE-2 agonists and antagonists and related DNA - useful for  
 PT promoting or blocking neovascularisation, etc

XX Claim 15; Fig 6; 113pp; English.

|    |  |               |
|----|--|---------------|
| PD | 18-APR-1996.   |               |
| XX |  |               |
| PF | 06-OCT-1995;   | 95WO-US12935. |
| XX |  |               |
| PR | 06-APR-1995;   | 95US-0418595. |
| XX |  |               |
| PR | 27-OCT-1994;   | 94US-0319932. |
| PR | 27-OCT-1994;   | 94US-0330261. |
| PR | 02-DEC-1994;   | 94US-0348492. |
| PR | 09-DEC-1994;   | 94US-0353503. |
| PR | 17-JAN-1995;   | 95US-0373579. |
| XX |  |               |
| PA | (REGE-) REGENERON PHARM INC.   |               |
| XX |  |               |
| PI | Aldrich TH, Bruno J, Davis S, Goldfarb M, Jones PF,                      |               |
| PI | Maisonpierre PC, Radziejewski C, Yancopoulos GD;                         |               |
| PI | WPI: 1996-209850/21.   |               |
| DR | N-PSDB: AAT14650.  |               |
| DR |  |               |
| XX |  |               |
| PT | Nucleic acid encoding TIE-2 ligand and related vectors - useful in       |               |
| PT | diagnosis and treatment of neovascularisation, tumours, etc., or to      |               |
| PT | promote wound healing, etc.  |               |
| XX |  |               |
| XX | Claim 2; Fig 6; 84pp; English.   |               |
| PS |  |               |
| XX | AAR94605 is a human TIE-2 (HTIE-2) ligand 2 derived from a pBluescript   |               |
| CC | KS clone. HTIE-2 ligand DNAs of the invention are recombinant versions   |               |
| CC | of the native ligand coding sequences and may be used to produce the     |               |
| CC | ligands at a high yield. Antibodies and receptor bodies that bind to     |               |
| CC | TIE-2 ligands may be used to inhibit angiogenesis and neovascularisation |               |
| CC | (e.g. associated with tumour development) and the TIE-2 ligands          |               |
| CC | themselves are useful to promote neovascularisation and wound healing    |               |
| CC | e.g. for treatment of ischaemia. TIE-2 ligands are also useful to        |               |
| CC | treat thromboembolytic disease, atherosclerosis, inflammation and        |               |
| CC | diabetes. Ligand bodies contg. TIE-2 ligands may also be useful for      |               |
| CC | the delivery and targeting of growth factors, toxins etc. to sites       |               |
| CC | where their presence is advantageous.                                    |               |
| XX |  |               |
| SQ | Sequence 496 AA:   |               |
| XX |  |               |
| XX | Query Match 22.0%: Score 522; DB 17; Length 496;                         |               |
| XX | Best Local Similarity 27.9%: Pred. No. 4.4e-38;                          |               |
| XX | Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 1         |               |
| QY | 43 LESRGKE---EAGECPYVSLPPL-----TIQ---LPKQFSRIEEVFKEVQN 84                |               |
| DB | :::   : :  |               |
| DB | 26 MDSTGKKQYQVGHGSCSYTFLLPMDNCSSSPSYVNAVQDAPLEYDSDVQRVLQVLE 85           |               |
| QY | 85 LKE-----IVNSLKKSCDCKLQADDNGDPGRNGLLPSTGAPGEVDN--- 128                 |               |
| DB | :   :  |               |
| DB | 86 IMENNTQWLKLENYIQDNMKKEVYEQNAVQN-----QTAVMIEIGTNLLNQ 134               |               |
| QY | 129 -----RVRELESEV-----NKLSELKNKKEINVNLGR---LEKL 161                     |               |
| DB | ::: :::::  |               |
| DB | 135 TAEQTRKLTIDVEAQVNLQTTRELQLLEHSLSTNKKLEKLDIQTOSTINKLDKNSFLEKK 194     |               |
| QY | 162 NLV-----NNNNIENYVDSKANLTF-----V 183                                  |               |
| DB | :  |               |
| DB | 195 VLAMEDKHIIQLSTKEBKDQLVLYSKNSIIELEKIVTATVNNVSLQKQHDLMET 254           |               |
| QY | 184 VNSLDGKCKSPSQIQSRPVQHLKYKCDSDYYAIGKRSSEYRVTPDPKNSSEFYVCD 243         |               |
| DB | :  |               |
| DB | 255 VNNLLTMMSTNSAKDPTVAKEQIISFRCAEVKPSGHTTNGIYTLTFPNSTEBIKAYCD 314       |               |
| QY | 244 METMGSGWTVLQARLDGTSNFTFTQDYKAGFGNLRREFWLGNDKTHLLTKSKEMILRID 303      |               |
| DB | :  |               |
| DB | 315 MEAGGGWTIIQRREDGSDVFQRTWKKEYKVGFGNPSGEYWLGNFVQSOLTNQORYVLKIH 374     |               |
| QY | 304 LEDPENGVELYALDQFYVANEFLKYLPHLVGNYNGTAGDALRFNKHYNHDLKFTTPDKON 363     |               |
| DB | :  |               |
| DB | 375 LKDWEGNEAYSLYHFLYLSSEELNYRIHLGUTGATKISSISQPGND----FSTKDGDN 430       |               |
| QY | 364 DRYPGNCGLYYSSGWFWDACLANSNLNGKYYHQYRGVR-NGIFWGTWPGVSEAHPGYK 422       |               |

```

Db 431 DKIC-KCSOMUTGGMWFDACPSNLNGMYTPQONTNKFNGIKWYWG-----SGY- 482
QY 423 SSFKEAKMMIRPKHF 437
Db 483 -SLKATTMMIRPADF 496

RESULT 6
AAW47532
ID AAW47532 standard; Protein; 496 AA.
XX
AC AAW47532;
DT 09-SEP-1998 (first entry)
DE Human TIE-2 ligand 2 from clone pBluescript KS.
XX Chimeric TIE ligand 2N1CF; TIE-2 ligand; neovascularisation;
KW tumour; human.
KW Homo sapiens.
PN WO9805779-A1.
XX
PD 12-FEB-1998.
XX
PF 01-AUG-1997; 97WO-US13557.
XX
PR 25-OCT-1996; 96US-0740223.
PR 02-AUG-1996; 96US-0022999.
XX
PA (RECE-) REGENERON PHARM INC.
XX
PI Davis S, Yancopoulos GD;
XX WPI; 1998-145615/13.
DR N-PSDB; AAV18619.
XX
PT Modified human TIE-2 receptor ligand(s) - useful for promoting wound
healing
XX
PS Example 8; Fig 6; 202pp; English.
XX
CC This is the amino acid sequence of the human TIE-2 ligand 2,
used in the method of the invention, involving the production
of TIE-2 ligands which promote healing. The nucleic acids, vectors
and host cells used in the method of the invention are useful for
the recombinant production of the ligands. The ligands, etc. are
useful for blocking blood vessel growth, promoting neovascularisation,
promoting the growth or differentiation of a cell expressing the TIE
receptor, blocking the growth or differentiation of a cell expressing
the TIE receptor and for attenuating or preventing tumour growth in
a human.
XX
SQ Sequence 496 AA;

Query Match 22.0%; Score 522; DB 19; Length 496;
Best Local Similarity 27.9%; Pred. No. 4.4e-38;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCF--EAGECPYQVSLPPL-----TIQ--LPKQFSRIEVEFKEVON 84
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 26 MDSICKKQYQVHGSCSYFLPDPENDNCRSSSPYVSNVAVQADAPLEVDSVQRLQVLEN 85
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 85 LKE-----IVNSLKKSCODCKLQADNDGPGRGNLLPSTGACGEVGDN---- 128
:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 86 IMENNTWMLKLENYTODNNKKEWVETQNAVQN-----QAVMEIGTINLQ 134
:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 129 -----RVRELESEV-----NKLSSLLKNAKEEINVLHGR---LEKL 161
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 135 TAEQTRKLTDEAQLNQVLTQRLLEHLSLSTNKLKQILDQTSINKLQDNKSFLEKK 194
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

QY 162 NLV-----NNNNIENVDSKVANLTF-----V 183
Db 195 VLAMEDKHIIQLSKEKQLOLVIVSKONSTIELEKKIVTATVNNNSVLQKQHDLMET 254
QY 184 VNSLDGKCKSCPSQEQIQSRPVQHLYKDCSDYYAIGKRSSSEYRTVTPDPKNSFEVYCD 243
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 255 VNNLLTMMSTNSKADPTVAKEEQISFRDCAEVFKSGHTTNGIYTUTFPNSTEEIKAYCD 314
QY 244 METMGGGWTVLQARLDGSTNFTTWQDYKAGFNLNREFWLGNDKTHLLTKSKMILRID 303
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 315 MEAGGGGWTVIQRREDGSDVDFQRTWKEYKVGPNSGEYVLGNEFVSQLTNOQRYVLKIH 374
QY 304 LEDFNGVELYALYDOFYVANEFLKYLHVGNVNGTAGDALRFNKHYNHDLKFFFTPDKN 363
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 375 LKDWEGNEAYSLEYEHFYLSSSEELNYRIHLKGLTAGKISSIQPGND----FSTKGDGN 430
QY 364 DRYPSGNCGLYSSGWEDACLSANLNGKYYHQYRGVR-NGIFMGCTWPGVSEAHPGGYK 422
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 431 DKIC-KCSOMUTGGMWFDACPSNLNGMYTPQONTNKFNGIKWYWG-----SGY- 482
QY 423 SSFKEAKMMIRPKHF 437
Db 483 -SLKATTMMIRPADF 496

RESULT 7
AAB28392
ID AAB28392 standard; Protein; 496 AA.
XX
AC AAB28392;
DT 19-FEB-2001 (first entry)
DE Human angiopoietin-2.
XX
KW Human; angiopoietin-2; cytostatic; antiproliferative;
KW vascular endothelial growth factor; VEGF; antibody; VEGF2 receptor;
KW blood vessel regression; cancer; vascularised solid tumour.
XX
OS Homo sapiens.
XX
PN WO2000064946-A2.
XX
PD 02-NOV-2000.
XX
PF 28-APR-2000; 2000WO-US11367.
XX
PR 28-APR-1999; 99US-0131432.
XX (TEXA) UNIV TEXAS SYSTEM.
XX PI Thorpe PE, Brekken RA;
XX WPI; 2000-687317/67.
DR N-PSDB; AAC67774.
XX
PT Immunogenic composition for the treatment and diagnosis of cancer
comprises an anti-VEGF (vascular endothelial growth factor) antibody
binding the same epitope as the monoclonal antibody ATCC PTA 1595 -
XX
PS Disclosure; Page 283-285; 298pp; English.
XX
CC The present invention relates to anti-Vascular Endothelial Growth Factor
(VEGF) antibodies that bind to the same epitope as the monoclonal
antibody ATCC PTA 1595 and which significantly inhibit VEGF binding to
the VEGF receptor VEGFR2, without inhibiting VEGF binding to the VEGF
receptor VEGFR1. The present sequence is human angiopoietin-2.
CC Angiopoietin-2 may be operatively attached to the anti-VEGF antibodies of
the present invention. Angiopoietin-2 acts to disturb capillary structure
and is thought to lead to vessel regression. The anti-VEGF antibodies of
the present invention are useful for the treatment and diagnosis of
cancer, especially vascularised solid tumours.
XX

```

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:01 11 November 2014







XX Modified human TIE-2 receptor ligand(s) - useful for promoting wound  
 PT healing  
 XX  
 XX  
 PS Claim 20; Fig 24; 202pp; English.  
 XX  
 CC This is the amino acid sequence of the chimeric TIE ligand IN1C2F,  
 CC used in the method of the invention, involving the production  
 CC of TIE-2 ligands which promote healing. The nucleic acids, vectors  
 CC and host cells used in the method of the invention are useful for  
 CC the recombinant production of the ligands. The ligands, etc. are  
 CC useful for blocking blood vessel growth, promoting neovascularisation,  
 CC promoting the growth or differentiation of a cell expressing the TIE  
 CC receptor, blocking the growth or differentiation of a cell expressing  
 CC the TIE receptor and for attenuating or preventing tumour growth in  
 CC a human.  
 XX  
 SQ Sequence 498 AA;

Query Match 21.9%; Score 521.5; DB I9; Length 498;  
 Best Local Similarity 32.7%; Pred. No. 5e-38;  
 Matches 131; Conservative 61; Mismatches 138; Indels 71; Gaps 14;  
 QY 61 LPPLTIQLPKQFSRIE-----EVEKEVQNLKEIVNSLKKSCQDCKLQADDNGDPGRN 112  
 DB 145 LTDVETQVLNQTSLRIQLLENSLSTYKLEKQLQQTWEILKHE-----KN 191  
 QY 113 GLLPSTGAPGCVGNRVRLESEVNKLSSELKNAKEEINVHLC-----RLEK-L 161  
 DB 192 SLL-----EHKLEWE---GKHKEELDTKEKENLQGLVTRQTYITQIELEKQL 237  
 QY 162 NLVMNNIENVDSKANLTFVNSLDGKCSK---CPSQIQSRPVQHLIYKDCSDY 217  
 DB 238 NRATTNN--SVLQKQLESLMDTVHNLCTKEGVLLKGGKREEKP-----FRDCAEVE 290  
 QY 218 AIGKRSETVRYTPDPKNSFEVYCDMETMGGGTWVLOARLDGSTNFTTWQDYKAGFGN 277  
 DB 291 KSGHTNGIYTLTFNSTEEIKAYCDEAGGGGTIIQRREDGSDVDFQRTWKEYKVGFGN 350  
 QY 278 LRREFWGNDKIHLLTKSKEMILRIDLEDFNGVELYALYDOFYVANEFLKYRLHVGNYNG 337  
 DB 351 PSGEYELWEGNEFVSQLTNOQRYVLKIHLDKOWEGNEAYSLYEDFVLSSEELNRYRIHLKGLTG 410  
 QY 338 TAGDALRFNKHNDLKFTTTPDKNDRYPSGNCGLIYSSQWDFDACL SANLNGKYHOKYRGVR 397  
 DB 411 TAGKISSISOPND-----FSTKGDNDKIC--KCSQMLTGWDFDAGPSNLNGMYYPQR 465  
 QY 398 YRGVR-NGIFWGTWPGVSEAHPGGYKSSFKKAKMMIRPKHF 437  
 DB 466 QNTNKGNGIKWYWRG-----SGY--SLKATTMMIRPADF 498

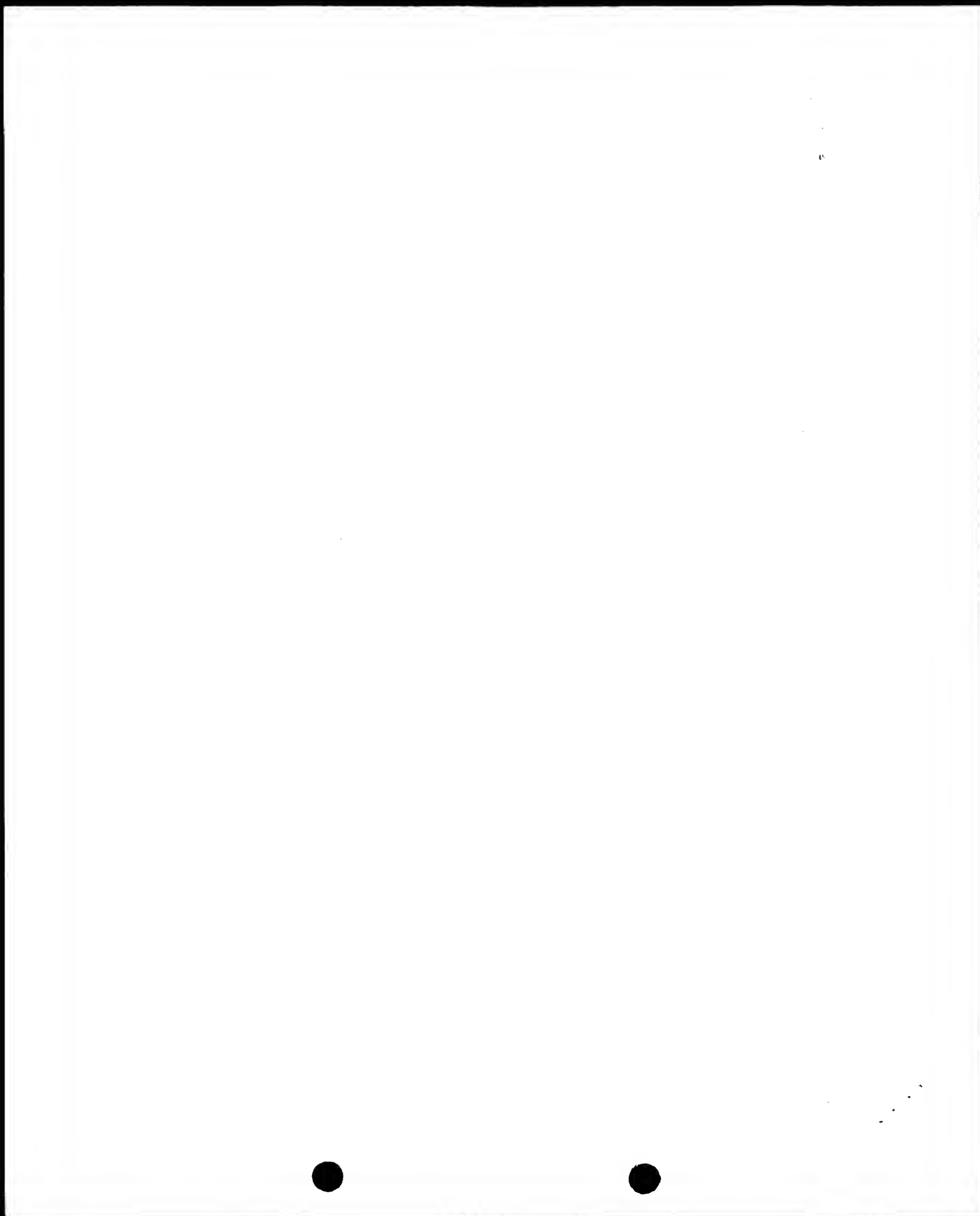
RESULT 14  
 AAU77252  
 ID AAU77252 standard; Protein; 444 AA.  
 XX  
 AC AAU77252;  
 XX  
 XX 20-MAY-2002 (first entry)  
 XX  
 XX Human angiotensin-2-443 protein.  
 XX  
 XX Human; angiotensin-2-443; Ang2\_443; Tie-2 receptor-related disorder.  
 XX  
 XX Homo sapiens.  
 XX  
 XX KR2001070139-A.  
 XX  
 XX 25-JUL-2001.  
 XX  
 XX 13-OCT-2000; 2000KR-0060202.  
 XX  
 XX 14-OCT-1999; 99KR-004605.  
 XX  
 PR

XX (BIOR-) BIORROUTE CO LTD.  
 XX  
 XX Kim IJ, Ko GY;  
 XX  
 DR WPI; 2002-065267/09.  
 DR N-PSDB; ABK10923.  
 XX  
 XX Human angiotensin-2-443 protein and its gene -  
 XX  
 PS Disclosure; Page 12; 17pp; Korean.  
 XX  
 CC The present invention relates to the isolation of human  
 CC angiotensin-2-443 (Ang2\_443) protein and the gene encoding it.  
 CC The sequences of the invention can be used in the prevention and  
 CC treatment of diseases related to cells which express the Tie-2 receptor.  
 CC The present sequence represents human Ang2\_443 protein.  
 XX  
 SQ Sequence 444 AA;

Query Match 21.9%; Score 520; DB 23; Length 444;  
 Best Local Similarity 29.6%; Pred. No. 5.7e-38;  
 Matches 134; Conservative 74; Mismatches 154; Indels 90; Gaps 15;  
 QY 43 LESRGKCE--EAGCEPYQVSLPL-----TIDLPKQFSRIE-----EVEKEVQNLKEIVN 90  
 DB 26 MDSIGKKQYQVHSGSCSTFLFEMONCRSSSPVSVNAVORDAPLEYDDSVQRLQVLEN 85  
 QY 91 SLKKSC-----QDCKLQADDNGDPGRNGLLPSTGAPGCVGNRVRLESEVNKLS 142  
 DB 86 IMENNTQWMLKVNQTRLELQ-----LLEHSLSTNKL-EKQILDQTSINKLQD 134  
 QY 143 -----ELKNAKEEINVHLCRLKLEKLVNMMNNTIENVYDSKVAN----- 179  
 DB 135 KNSFEKKVLAMEDKHIIQIQLQSTKEEKQQLQVLYSKNSI-IEELEKKIVTATVNNVLQ 193  
 QY 180 -----LTFVNSLDGKCSKPSQIQSRPVQHLIYKDCSDYVAGKRSET 226  
 DB 194 KOQHLDMETVNNLLTMTMSTNSAKDPTVAKEQIS-----FRDCAEVEKSGHTNGI 245  
 QY 227 YRVTPDPKNSFEVYCDMETMGGGTWVLOARLDGSTNFTTWQDYKAGFGNLRREFWGN 286  
 DB 246 YTLTFNSTEEIKAYCDEAGGGGTIIQRREDGSDVDFQRTWKEYKVGFGNPSGEYELGN 305  
 QY 287 DKHLTKSKEMILRIDLEDFNGVELYALYDOFYVANEFLKYRLHVGNYNGTAGDALRFN 346  
 DB 306 EFVSQLTNOQRYVLKIHLDKOWEGNEAYSLYEHFVLSSEELNRYRIHLKGLTGAGKISSIS 365  
 QY 347 KHYNHDLKFTTTPDKNDRYPSGNCGLIYSSQWDFDACL SANLNGKYHOKYRGVR-NGI 405  
 DB 366 QPGND-----FSTKGDNDKIC--KCSQMLTGWDFDAGPSNLNGMYYPQRONTNKFNGI 420  
 QY 406 FWGTWPGVSEAHPGGYKSSFKKAKMMIRPKHF 437  
 DB 421 KWIYWRG-----SGY--SLKATTMMIRPADF 444

RESULT 15  
 AAU77529  
 ID AAU77529 standard; Protein; 499 AA.  
 XX  
 AC AAU77529;  
 XX  
 XX 09-SEP-1998 (first entry)  
 XX  
 XX Amino acid sequence of chimeric TIE ligand IN2C2F (chimeric 3).  
 XX  
 XX Chimeric TIE ligand IN2C2F; TIE-2 ligand; neovascularisation;  
 KW tumour; human.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO9805779-A1.  
 XX  
 PN





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:53:36 ; Search time 13.5374 Seconds  
(without alignments)  
516.399 Million cell updates/sec

Title: US-09-902-563-2

Perfect score: 2378

Sequence: 1 MKLANWYWLSSAVLATYGLF.....GYKSSFKAKMIRPKHFKP 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                 | Description       |
|------------|-------|-------------|--------|-----------------------|-------------------|
| 1          | 522   | 22.0        | 496    | 9 US-10-179-744-6     | Sequence 6, Appli |
| 2          | 522   | 22.0        | 496    | 10 US-09-998-831-4    | Sequence 4, Appli |
| 3          | 522   | 22.0        | 496    | 10 US-09-997-306-14   | Sequence 14, Appl |
| 4          | 506   | 21.3        | 339    | 10 US-09-925-301-1082 | Sequence 1082, Ap |
| 5          | 505   | 21.2        | 491    | 10 US-09-897-306-1    | Sequence 1, Appli |
| 6          | 487   | 20.5        | 497    | 9 US-10-179-744-4     | Sequence 4, Appli |
| 7          | 486.5 | 20.5        | 498    | 10 US-09-998-831-2    | Sequence 2, Appli |
| 8          | 486.5 | 20.5        | 498    | 10 US-09-897-306-13   | Sequence 13, Appl |
| 9          | 480.5 | 20.2        | 493    | 10 US-09-818-143-23   | Sequence 23, Appl |
| 10         | 480.5 | 20.2        | 498    | 9 US-10-179-744-2     | Sequence 2, Appli |
| 11         | 478.5 | 20.1        | 495    | 10 US-09-998-831-5    | Sequence 5, Appli |
| 12         | 468   | 19.7        | 411    | 9 US-09-912-740A-1    | Sequence 1, Appli |
| 13         | 468   | 19.7        | 411    | 10 US-09-912-741B-1   | Sequence 1, Appli |
| 14         | 460.5 | 19.4        | 461    | 9 US-09-992-598-314   | Sequence 314, App |
| 15         | 460.5 | 19.4        | 461    | 10 US-09-989-722-314  | Sequence 314, App |
| 16         | 460.5 | 19.4        | 461    | 10 US-09-989-723-314  | Sequence 314, App |
| 17         | 460.5 | 19.4        | 461    | 10 US-09-989-279-314  | Sequence 314, App |
| 18         | 460.5 | 19.4        | 461    | 10 US-09-989-727-314  | Sequence 314, App |
| 19         | 460.5 | 19.4        | 461    | 10 US-09-989-731-314  | Sequence 314, App |

|    |       |      |      |                        |                   |
|----|-------|------|------|------------------------|-------------------|
| 20 | 460.5 | 19.4 | 461  | 10 US-09-989-732-314   | Sequence 314, App |
| 21 | 460.5 | 19.4 | 461  | 10 US-09-991-073-314   | Sequence 314, App |
| 22 | 460.5 | 19.4 | 461  | 10 US-09-990-442-314   | Sequence 314, App |
| 23 | 460.5 | 19.4 | 461  | 10 US-09-991-163-314   | Sequence 314, App |
| 24 | 460.5 | 19.4 | 461  | 10 US-09-993-604-314   | Sequence 314, App |
| 25 | 460.5 | 19.4 | 461  | 10 US-09-990-456-314   | Sequence 314, App |
| 26 | 460.5 | 19.4 | 461  | 10 US-09-989-721-314   | Sequence 314, App |
| 27 | 451   | 19.0 | 847  | 9 US-10-112-527-4      | Sequence 4, Appli |
| 28 | 449   | 18.9 | 251  | 9 US-10-112-527-1      | Sequence 1, Appli |
| 29 | 443.5 | 18.7 | 269  | 9 US-09-966-546-26     | Sequence 26, Appl |
| 30 | 443.5 | 18.7 | 269  | 9 US-09-966-545-26     | Sequence 2, Appli |
| 31 | 438.5 | 18.4 | 236  | 9 US-10-112-527-2      | Sequence 2, Appli |
| 32 | 438.5 | 18.4 | 269  | 9 US-09-966-546-28     | Sequence 28, Appl |
| 33 | 438.5 | 18.4 | 269  | 9 US-09-966-545-28     | Sequence 28, Appl |
| 34 | 434.5 | 18.3 | 221  | 9 US-10-112-527-3      | Sequence 3, Appli |
| 35 | 388   | 16.3 | 2167 | 10 US-09-778-927A-61   | Sequence 61, Appl |
| 36 | 388   | 16.3 | 2201 | 12 US-10-100-912-2     | Sequence 2, Appli |
| 37 | 363.5 | 15.3 | 405  | 9 US-10-000-512-6      | Sequence 6, Appli |
| 38 | 304   | 12.8 | 59   | 10 US-09-864-761-41396 | Sequence 41396, A |
| 39 | 291   | 12.2 | 510  | 10 US-09-925-301-1131  | Sequence 1131, Ap |
| 40 | 232.5 | 9.8  | 489  | 9 US-09-989-920-185    | Sequence 185, App |
| 41 | 147.5 | 6.2  | 133  | 10 US-09-925-297-858   | Sequence 858, App |
| 42 | 125   | 5.3  | 55   | 10 US-09-864-761-39830 | Sequence 39830, A |
| 43 | 114.5 | 4.8  | 480  | 9 US-09-859-888-6      | Sequence 6, Appli |
| 44 | 109.5 | 4.6  | 480  | 9 US-09-859-888-5      | Sequence 5, Appli |
| 45 | 108   | 4.5  | 87   | 10 US-09-939-825-29    | Sequence 29, Appl |

#### ALIGNMENTS

RESULT 1  
US-10-179-744-6  
; Sequence 6, Application US/10179744  
; Patent No. US20020173627A1  
; GENERAL INFORMATION:  
; APPLICANT: Davis, Samuel et al.  
; TITLE OF INVENTION: TIE-2 LIGANDS, METHODS OF MAKING AND USES THEREOF  
; FILE REFERENCE: REG 330-F-PCT-US  
; CURRENT APPLICATION NUMBER: US/10/179,744  
; CURRENT FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: US/08/817,318  
; PRIOR FILING DATE: 1995-09-16  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Unknown Organism  
US-10-179-744-6

|                       |                  |  |             |             |
|-----------------------|------------------|--|-------------|-------------|
| Query Match           | 22.0%            | Score 522;   | DB 9;       | Length 496; |
| Best Local Similarity | 27.9%;           | Pred. No. 7e-35;   |             |             |
| Matches 138;          | Conservative 74; | Mismatches 159;  | Indels 124; | Gaps 15;    |
| Qy                    | 43               | LESRGKCE---EAGECPYQVSVLPPL-----TIQ--LPKQFSRIEEVFEKVN 84          |             |             |
| Db                    | 26               | MDSTGKQYQVHGSCSYTFLLPNDNCSSSPYVNAVORDAPLEVDSDVQRLQVLEN 85        |             |             |
| Qy                    | 85               | LKE-----IVNSLKSCODCKLOADDGPGRNGLLLPSTGAPGEVDN---- 128            |             |             |
| Db                    | 86               | IMENNTOMAKLENYIQDNKKKEVETQNAVQN-----QTAVMEITGTLLN 134            |             |             |
| Qy                    | 129              | -----RVRELESEV-----NKLSSSELKNAKEEINVLHGR---LEKL 161              |             |             |
| Db                    | 135              | TAEQTRKLTIDEVQVNLQNTTRLEQLLHLSLSLKLEKQILDQTSINKLQDNKSFLEK 194    |             |             |
| Qy                    | 162              | NLV-----NMNNIYVDSKVANLTF-----V 183                               |             |             |
| Db                    | 195              | VLAMEDKHIIQLOSTKEEKDQLQVLVSKNSIIELEKKTIVTATVNNNSVLQKQOHDLMET 254 |             |             |

QY 184 VNSLDGKCKSPQEQIOQRVQHLIYKDCSDYYAIGKRSSTYRVTPDPKNSFEVYCD 243  
DB 255 VNNLTMMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTPNSTEETKAYCD 314  
QY 244 METGGGWTVLQARLDGSTNFTRWODYKAGFNLRRFELGNDKIHLLTKSKEMILRID 303  
DB 315 MEAGGGWTIIQRREDGSDVDFORTWKEYKVGFGNPSGEYWLGNFEVSQLTNQORYVLKIH 374  
QY 304 LEDFNGVELYALYDOFYVANEFLKYRLHVGNYGTAGDALRPNKHYNHDLKFTTTPDKDN 363  
DB 375 LKDWEGNEAYSLEYEHFYLSSELNRYHLKGLTGTAGKISSISOPGND----FSTKGDGN 430  
QY 364 DRYPSGNCGLIYSSGWDFDACL SANLNGKYHOKYRGVR-NGIFWGTWPGVSEAHPGGYK 422  
DB 431 DKCIC-KCSQMLTGGWDFDAGPSNLNGMYYPQONTNKGINKWYKWK-----SGY- 482  
QY 423 SSFKEAKMIRPKHF 437  
DB 483 -SLKATMMIRPADF 496  
RESULT 2  
US-09-998-831-4  
; Sequence 4, Application US/09998831  
; Patent No. US20020119153A1  
; GENERAL INFORMATION:  
; APPLICANT: Phillip E. Thorpe  
; APPLICANT: Rolf A. Brecken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY  
; FILE REFERENCE: 4001.002584  
; CURRENT APPLICATION NUMBER: US/09/998.831  
; PRIOR FILING DATE: 2001-11-30  
; PRIOR FILING DATE: 2001-11-30  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-998-831-4  
Query Match 22.0%; Score 522; DB 10; Length 496;  
Best Local Similarity 27.9%; Pred. No. 7e-35;  
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVON 84  
DB 26 MDSIGKKQYQVHGSCSYTFLLPEMDCNRSSSPYVSNVQADAPLEYDDSVQRLQVLEN 85  
QY 85 LKE-----IVNSLKKSCODCKLQADNGDPGRNGLLLPSTGAPGEVGDN---- 128  
DB 86 IMENNTOMLKMENYIODNNKKEVIEIQONAVQN-----QTAVMIEIGTNLLNQ 134  
QY 129 -----RVRELESEV-----NKLSELKNAKEEINVLHGR---LEKL 161  
DB 135 TABQTRKLTDEAQLVNOTRLELQLEHLSLSTNKLKQILQDQSEINKLQDNSELEKK 194  
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183  
DB 195 VLAMEDKHIIQLOSIKEEKDQQLVLSKQNSIIELEKKIVTATVNNSVLQKQOHDLMET 254  
QY 184 VNSLDGKCKSPQEQIOQRVQHLIYKDCSDYYAIGKRSSTYRVTPDPKNSFEVYCD 243  
DB 255 VNNLTMMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTPNSTEETKAYCD 314  
QY 244 METGGGWTVLQARLDGSTNFTRWODYKAGFNLRRFELGNDKIHLLTKSKEMILRID 303  
DB 315 MEAGGGWTIIQRREDGSDVDFORTWKEYKVGFGNPSGEYWLGNFEVSQLTNQORYVLKIH 374  
QY 304 LEDFNGVELYALYDOFYVANEFLKYRLHVGNYGTAGDALRPNKHYNHDLKFTTTPDKDN 363  
DB 375 LKDWEGNEAYSLEYEHFYLSSELNRYHLKGLTGTAGKISSISOPGND----FSTKGDGN 430  
QY 364 DRYPSGNCGLIYSSGWDFDACL SANLNGKYHOKYRGVR-NGIFWGTWPGVSEAHPGGYK 422  
DB 431 DKCIC-KCSQMLTGGWDFDAGPSNLNGMYYPQONTNKGINKWYKWK-----SGY- 482  
QY 423 SSFKEAKMIRPKHF 437  
DB 483 -SLKATMMIRPADF 496

DB 375 LKDWEGNEAYSLEYEHFYLSSELNRYHLKGLTGTAGKISSISOPGND----FSTKGDGN 430  
QY 364 DRYPSGNCGLIYSSGWDFDACL SANLNGKYHOKYRGVR-NGIFWGTWPGVSEAHPGGYK 422  
DB 431 DKCIC-KCSQMLTGGWDFDAGPSNLNGMYYPQONTNKGINKWYKWK-----SGY- 482  
QY 423 SSFKEAKMIRPKHF 437  
DB 483 -SLKATMMIRPADF 496  
RESULT 3  
US-09-897-306-14  
; Sequence 14, Application US/09897306  
; Patent No. US20020123054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Gorgone, Gina A.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Murry, Lynn E.  
; TITLE OF INVENTION: HUMAN ANGIOPOIETIN  
; FILE REFERENCE: PC-0048.CIP  
; CURRENT APPLICATION NUMBER: US/09/897.306  
; CURRENT FILING DATE: 2001-07-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PERL Program  
; SEQ ID NO 14  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20020123054A1 g2257933  
US-09-897-306-14  
Query Match 22.0%; Score 522; DB 10; Length 496;  
Best Local Similarity 27.9%; Pred. No. 7e-35;  
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGECPYQVSLPPL-----TIQ--LPKQFSRIEEVFKEVON 84  
DB 26 MDSIGKKQYQVHGSCSYTFLLPEMDCNRSSSPYVSNVQADAPLEYDDSVQRLQVLEN 85  
QY 85 LKE-----IVNSLKKSCODCKLQADNGDPGRNGLLLPSTGAPGEVGDN---- 128  
DB 86 IMENNTOMLKMENYIODNNKKEVIEIQONAVQN-----QTAVMIEIGTNLLNQ 134  
QY 129 -----RVRELESEV-----NKLSELKNAKEEINVLHGR---LEKL 161  
DB 135 TABQTRKLTDEAQLVNOTRLELQLEHLSLSTNKLKQILQDQSEINKLQDNSELEKK 194  
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183  
DB 195 VLAMEDKHIIQLOSIKEEKDQQLVLSKQNSIIELEKKIVTATVNNSVLQKQOHDLMET 254  
QY 184 VNSLDGKCKSPQEQIOQRVQHLIYKDCSDYYAIGKRSSTYRVTPDPKNSFEVYCD 243  
DB 255 VNNLTMMSTNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTPNSTEETKAYCD 314  
QY 244 METGGGWTVLQARLDGSTNFTRWODYKAGFNLRRFELGNDKIHLLTKSKEMILRID 303  
DB 315 MEAGGGWTIIQRREDGSDVDFORTWKEYKVGFGNPSGEYWLGNFEVSQLTNQORYVLKIH 374  
QY 304 LEDFNGVELYALYDOFYVANEFLKYRLHVGNYGTAGDALRPNKHYNHDLKFTTTPDKDN 363  
DB 375 LKDWEGNEAYSLEYEHFYLSSELNRYHLKGLTGTAGKISSISOPGND----FSTKGDGN 430  
QY 364 DRYPSGNCGLIYSSGWDFDACL SANLNGKYHOKYRGVR-NGIFWGTWPGVSEAHPGGYK 422  
DB 431 DKCIC-KCSQMLTGGWDFDAGPSNLNGMYYPQONTNKGINKWYKWK-----SGY- 482  
QY 423 SSFKEAKMIRPKHF 437  
DB 483 -SLKATMMIRPADF 496

Dd 483 -SLKATTMMIRPADE 496

## RESULT 4

US-09-925-301-1082  
; Sequence 1082, Application US/09925301  
; Patent No. US20020052308A1

## GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1082  
LENGTH: 339  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (42)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-301-1082

Query Match 21.3%; Score 506; DB 10; Length 339;  
Best Local Similarity 44.6%; Pred. No. 8.3e-34;  
Matches 107; Conservative 25; Mismatches 88; Indels 20; Gaps 6;

QY 210 YKDCSDYATGKRSSEYRVTDPKNSFEVYCDMETMGGWTLQARLDGSTNETRTWQ 269  
Dd 107 YADCEIFNDGYKLSFYKIRPLASPAEFVSVCDSMD-GGGWTVIQRSDGSENFNRGWK 165  
QY 270 DYKAGFGL---RREFWLGNDKIHLLTKSKEMILRIDLEDFNGVELYALDQPYVANEF 326  
Dd 166 DYENGFGNFVQKHGEYWLGNKHLHFLITQEDYTLKIDLADFEKNSRYAQKFNKVGDEKN 225  
QY 327 KYRLHVGNYNTAGDALREKNH-----YNHDLKFTPTDPKNDRYPSGNGLYYSSGW 380  
Dd 226 FYELNIGEYSTAGDSLACNFHPEVQWASHQWKFEIWDHDNY-EGNCAEEDDSGNW 284  
QY 381 FDACLSANLNGKYYHKGKRG-VRNGIFWGTWPGVSEAHPPGKYSFKEAKMTIRPKHEKP 439  
Dd 285 FNRCHSANLNGVYSGPYTAKTDNGIVWTW-----HGWYSLKSVVVMKIRPNDETP 336

## RESULT 5

US-09-897-306-1  
; Sequence 1, Application US/09897306  
; Patent No. US20020123054A1

## GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Gorgone, Gina A.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Murry, Lynn E.  
; TITLE OF INVENTION: HUMAN ANGIOPOIETIN  
; FILE REFERENCE: PC-0048 CIP  
; CURRENT APPLICATION NUMBER: US/09/897,306  
; CURRENT FILING DATE: 2001-07-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
LENGTH: 491  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. US20020123054A1 2365223CD1

US-09-897-306-1

Query Match 21.2%; Score 505; DB 10; Length 491;  
Best Local Similarity 26.9%; Pred. No. 1.7e-33;  
Matches 131; Conservative 81; Mismatches 137; Indels 138; Gaps 15;

QY 47 GKCEAGECPYQVSLPPLTIOLP-----KQFSRIEVEF--KEVONLKEIVNSLUKSCQ 97  
Dd 42 GK-EAACKCAYTFLVPEQRIITGPICVNTKGQDASTIKDMITRMDLENLAKDVSQRKREID 100  
QY 98 DCKLOADNDGDPGRNGLLPSTGAPGEVDNRVRESEVNKLSSELKNAKEEINVHLGR 157  
Dd 101 VLQLVVDVGN-----IVNEVKLLKESRNWNSRVTLQMQ 136  
QY 158 L-----EKLNLVNMNNIENYV-----DSKANLTFVVS-----L 187  
Dd 137 LLHEIIRKRDNSLELSOLENKLNVVTTEMLKMATRYRELEVKYASLTDLVNNQSVMITLL 196  
QY 188 DGKSCKCPQEOIQSRP-----VOHL----- 208  
Dd 197 EEQCLRIFSRQDTHVSPPLVQVVPQHIPNSQOYTPGLLGGNEIOROPGYPRDLMPPDLA 256  
QY 209 -----IYKDCSDYIAGKRSSEYRVTDPKNSFEVYCDMETMG 249  
Dd 257 TSPTKSPFKIPPVTFINEGPFKQQAQKAGHSVSGIYMIKPENSNGPQLWCENSLDPG 316  
QY 250 GWTVQLARLDGSTNETRTWQDYKAGFNLRRERFWLGNDKIHLLTKSKEMILRIDLEDFNG 309  
Dd 317 GWTVIQRTDGSVNFERNWYKKGFGNDIGEYWLGLENIYILSNQDNKYLLELEDWSD 376  
QY 310 VELYALDQPYVANEFYRLHVGNYNTAGDALRPNKHYNHDLKFTPTDPKNDRYPSG 369  
Dd 377 KKVAEYSSRPLEPSESEFYRLRLGTQYQGNAGDSMMW-----HNGKQFTLDRDKDNY-AG 430  
QY 370 NCGLYYSSGWMFDCALSNLNGKYYH-QKYRGV-RNGIFWGTWPGVSEAHPPGKYSFKE 427  
Dd 431 NCAHFHKGHWYNNCAHSLNGLVWYRGHYRSHKHODGIFWAEYRG-----GSY--SLRA 482  
QY 428 AKMTIRP 434  
Dd 483 VQMMIKP 489

## RESULT 6

US-10-179-744-4  
; Sequence 4, Application US/10179744  
; Patent No. US20020173627A1

## GENERAL INFORMATION:

; APPLICANT: Davis, Samuel et al.  
; TITLE OF INVENTION: TIE-2 LIGANDS, METHODS OF MAKING AND USES THEREOF  
; FILE REFERENCE: REG 330-F-PCT-US  
; CURRENT APPLICATION NUMBER: US/10/179,744  
; CURRENT FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: US/08/817,318  
; PRIOR FILING DATE: 1999-09-16  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
LENGTH: 497  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Unknown Organism

US-10-179-744-4

Query Match 20.5%; Score 487; DB 9; Length 497;  
Best Local Similarity 32.4%; Pred. No. 4.9e-32;  
Matches 131; Conservative 59; Mismatches 136; Indels 78; Gaps 16;

QY 61 LPPLTIQLPKQFSRIE-----EVEKEVQNLKEIVNSLUKSCQCKLOADNDGDPGRN 112  
Dd 145 LVDVETVQNTSRLEIQLLENSLSTYKLEKQLQQTNETILKIH-----KN 191  
QY 113 GULLPSTGAPGEVDNRVRESEVNKLSSELKNAKEEINVHLG-----RLEK-L 161

Db 192 SLI-----EHLKLEME---GKHKEELDTLKEENLQGLVTRQTYIIQLEKQL 237  
QY 162 NLVNMNINIYVDSKANLTFFVNSLDGKCSK---CPSEQIOISRPVQHLYIKDCSDYYA 218  
Db 238 NRATTNN--SVLQKQOQLEMDTVHNLVNLCTKEVLLKGGKREDEKDP-----PRDCADYQ 290  
QY 219 IGRSSSETYRV---TPDPKNSPEVYCDMETMGGGWTVLQARLDGSTNFTRTWQDYKAG 274  
Db 291 AGFNKSGIYTIYINNPEPK-----KVFCDNDVNGGWTVYQHRDGLSDFQGWKEYKMG 346  
QY 275 FGNLRREFWLGNDKTHLLTSKEMILRIDLEDPNGVELYALYDOFYVANEFLKYLHVG 334  
Db 347 FGNPSGEYWLGNFEFAITSQRYMLRIELMDWEGNRAYSQYDRFHIGNEKQNYRLYLK 406  
QY 335 YNCTAGDALRFNKHYNHDLKFFTPDKDNDRYPSGNCGLYSSGWNFDACLSANLNGKY 393  
Db 407 HTGTAGKQSSLIH--GAD---FSTKDADNDNCMC--KCALMTGGWFDACGPNLNGMF 461  
QY 395 HQ-KYRGVRNGIFWGTWPGVSEAHFPGYKSSFKKAKMIRPKHF 437  
Db 462 YTAGQNHGKLGKIKWHYFKGPS-----YSLRSTTMMIRPLDF 498

## RESULT 7

US-09-998-831-2  
; Sequence 2, Application US/09998831  
; Patent No. US20020119153A1  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brecken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY  
; TITLE OF INVENTION: INHIBITING VEGF  
; FILE REFERENCE: 4001.002584  
; CURRENT APPLICATION NUMBER: US/09/998.831  
; PRIOR FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: 09/561,108  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-09-998-831-2

Query Match 20.5%; Score 486.5; DB 10; Length 498;  
Best Local Similarity 32.3%; Pred. No. 5.4e-32;  
Matches 131; Conservative 59; Mismatches 136; Indels 79; Gaps 16;

QY 61 LPPLTIQLPKQFSRIE-----EVFKEVONLKEIYVNSLKSCQDCKLQADNDGPRN 112  
Db 145 LTDVETQVNLQTSRLTQLLENSLSTYKLEKQLQQTNEILKHE-----KN 191  
QY 113 GLLPSTGAPCEVDGNRVRESEVKNLSSELKNKEEINVLHG-----RLEK-L 161  
Db 192 SLI-----EHLKLEME---GKHKEELDTLKEENLQGLVTRQTYIIQLEKQL 237  
QY 162 NLVNMNINIYVDSKANLTFFVNSLDGKCSK---CPSEQIOISRPVQHLYIKDCSDYY 217  
Db 238 NRATTNN--SVLQKQOQLEMDTVHNLVNLCTKEVLLKGGKREDEKDP-----PRDCADY 290  
QY 218 IGRSSSETYRV---TPDPKNSPEVYCDMETMGGGWTVLQARLDGSTNFTRTWQDYKA 273  
Db 291 AGFNKSGIYTIYINNPEPK-----KVFCDNDVNGGWTVYQHRDGLSDFQGWKEYKMG 346  
QY 274 FGNLRREFWLGNDKTHLLTSKEMILRIDLEDPNGVELYALYDOFYVANEFLKYLHVG 333  
Db 347 FGNPSGEYWLGNFEFAITSQRYMLRIELMDWEGNRAYSQYDRFHIGNEKQNYRLYLK 406  
QY 334 YNCTAGDALRFNKHYNHDLKFFTPDKDNDRYPSGNCGLYSSGWNFDACLSANLNGKY 393  
Db 407 HTGTAGKQSSLIH--GAD---FSTKDADNDNCMC--KCALMTGGWFDACGPNLNGMF 461

QY 394 YHQ-KYRGVRNGIFWGTWPGVSEAHFPGYKSSFKKAKMIRPKHF 437  
Db 462 YTAGQNHGKLGKIKWHYFKGPS-----YSLRSTTMMIRPLDF 498

## RESULT 8

US-09-897-306-13  
; Sequence 13, Application US/09897306  
; Patent No. US20020123054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Gorgone, Gina A.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Murry, Lynn E.  
; TITLE OF INVENTION: HUMAN ANGIOPOIETIN  
; FILE REFERENCE: PC-0048 CIP  
; CURRENT APPLICATION NUMBER: US/09/897.306  
; CURRENT FILING DATE: 2001-07-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PERL Program  
; SEQ ID NO 13  
; LENGTH: 498  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20020123054A1 g1907327  
US-09-897-306-13

Query Match 20.5%; Score 486.5; DB 10; Length 498;  
Best Local Similarity 32.3%; Pred. No. 5.4e-32;  
Matches 131; Conservative 59; Mismatches 136; Indels 79; Gaps 16;

QY 61 LPPLTIQLPKQFSRIE-----EVFKEVONLKEIYVNSLKSCQDCKLQADNDGPRN 112  
Db 145 LTDVETQVNLQTSRLTQLLENSLSTYKLEKQLQQTNEILKHE-----KN 191  
QY 113 GLLPSTGAPCEVDGNRVRESEVKNLSSELKNKEEINVLHG-----RLEK-L 161  
Db 192 SLI-----EHLKLEME---GKHKEELDTLKEENLQGLVTRQTYIIQLEKQL 237  
QY 162 NLVNMNINIYVDSKANLTFFVNSLDGKCSK---CPSEQIOISRPVQHLYIKDCSDYY 217  
Db 238 NRATTNN--SVLQKQOQLEMDTVHNLVNLCTKEVLLKGGKREDEKDP-----PRDCADY 290  
QY 218 IGRSSSETYRV---TPDPKNSPEVYCDMETMGGGWTVLQARLDGSTNFTRTWQDYKA 273  
Db 291 AGFNKSGIYTIYINNPEPK-----KVFCDNDVNGGWTVYQHRDGLSDFQGWKEYKMG 346  
QY 274 FGNLRREFWLGNDKTHLLTSKEMILRIDLEDPNGVELYALYDOFYVANEFLKYLHVG 333  
Db 347 FGNPSGEYWLGNFEFAITSQRYMLRIELMDWEGNRAYSQYDRFHIGNEKQNYRLYLK 406  
QY 334 YNCTAGDALRFNKHYNHDLKFFTPDKDNDRYPSGNCGLYSSGWNFDACLSANLNGKY 393  
Db 407 HTGTAGKQSSLIH--GAD---FSTKDADNDNCMC--KCALMTGGWFDACGPNLNGMF 461  
QY 394 YHQ-KYRGVRNGIFWGTWPGVSEAHFPGYKSSFKKAKMIRPKHF 437  
Db 462 YTAGQNHGKLGKIKWHYFKGPS-----YSLRSTTMMIRPLDF 498

## RESULT 9

US-09-818-143-23  
; Sequence 23, Application US/09818143  
; Patent No. US20020019000A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, Michael G.  
; APPLICANT: Volkumth, Wayne  
; APPLICANT: Klingler, Tod M.  
; TITLE OF INVENTION: POLYNUCLEOTIDES COEXPRESSED WITH MATRIX-REMODELING GENES  
; FILE REFERENCE: PB-0004 CIP



```
QY 332 VGNVNGTAGDALREKHYHDLKFTTPDKNDRYPSGNGLYSSGWFEDACLSANLNG 391
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 402 LKGGTGTAGKSSLLH-GAD---FSTKDADNDCMC-KCALMLTGSWFEDACGFSNLNG 456
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
QY 392 KYIHQ-KYRGVRNGFWGTWGVSEAHPGGKSSPKKAKMMIRPKHF 437
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 457 MEYTAGNHGKNGIKWHYKGPS-----YSLRSTTMMIRPLDF 495
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||

RESULT 12
US-09-912-740A-1
: Sequence 1, Application US/09912740A
: Patent No. US20020169280A1
: GENERAL INFORMATION:
: APPLICANT: Altieri, Dario C
: APPLICANT: Languino, Lucia R
: APPLICANT: Thornton, George B
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING
: TITLE OF INVENTION: ENDOTHELIAL CELL AND FIBRINOGEN MEDIATED INFLAMMATION
: FILE REFERENCE: 300.1DIV3
: CURRENT APPLICATION NUMBER: US/09/912,740A
: PRIOR FILING DATE: 2002-05-07
: PRIOR APPLICATION NUMBER: US 09/347,877
: PRIOR FILING DATE: 1999-07-06
: PRIOR APPLICATION NUMBER: US 08/748,150
: PRIOR FILING DATE: 1996-11-12
: PRIOR APPLICATION NUMBER: US 08/232,532
: PRIOR FILING DATE: 1994-04-25
: PRIOR APPLICATION NUMBER: US 08/139,562
: PRIOR FILING DATE: 1993-10-19
: PRIOR APPLICATION NUMBER: US 07/898,117
: PRIOR FILING DATE: 1992-06-12
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 411
: TYPE: PRT
: ORGANISM: Unknown
: FEATURE:
: OTHER INFORMATION: expressed
: FEATURE:
: NAME/KEY: CARBOHYD
: LOCATION: 88
: OTHER INFORMATION: site of glycosylation
: FEATURE:
: NAME/KEY: DISULFID
: LOCATION: (153)...(182)
: OTHER INFORMATION: disulfide-bond
: FEATURE:
: NAME/KEY: DISULFID
: LOCATION: (326)...(339)
: OTHER INFORMATION: disulfide-bond
US-09-912-740A-1

Query Match 19.7%; Score 468; DB 9; Length 411;
Best Local Similarity 30.6%; Pred. No. 1.3e-30;
Matches 126; Conservative 54; Mismatches 168; Indels 64; Gaps 13;

QY 22 VANNETEEIKDERAKDVCYVRLSRGKCEEGECYQVSLPPLTIQLPKQFSRTEEVFKE 81
|| : ||||| : : ||||| : : ||||| : : |||||
Db 2 VATRDNCCILDERGSGYCP-----TCGIADFLSTYQTKVDKDLQSLIEDLHQ 49
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
QY 82 VONLKETVNSLKKSCQDCKLQADNDGPGRNGLLPSTGAPGEVGDNRVRESEVKNLS 141
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 50 VENKTSEVKQLIKA-----IOLTYPNDESSKPNMIDAATLKSRLMEELMKYEASILTLD 104
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
QY 142 SELKNAKEEINVLHGRLEKLNLMNNIENYVDSKVANLTFVYVNSLDGCKSCQFSQIQ 201
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 105 SSRIYLOE-----IYNSNN-----OKIYNLKEKVAQLEAQOE-PCKDTVQ 144
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
QY 202 SRPVQHLIYKDCSDYAIAGKRSSETYRVTPDKNSSFYVCDMETMGGGWTVLQARLDGS 261
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 145 ---IHDITGKDCQDIANKAKQSGLYFIKPLKANOQFLVYCEIDGSGNGWIVFQKRLDGS 201
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
```

```
QY 262 TNETRTWODYKAGENL-----RREFWLGNDKIHLLTKSKEM--ILRLDEDNGVVELYAL 315
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 202 VDFKKNIQTKGEGFHLSPPTTTEFWLGNKEIHLISTQSALPVALRVLEEDWNGRTSTAD 261
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
QY 316 YDOFYVANEFKLYRLHGVN-NGTAGDAL-REKHYHDLKFFET-----TPDKONDR 365
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 262 YAMFKVGPADKYRLTYAYFAGGDAGDAFDGDFGDDPSDKFFFTSHNGMQFSTWMDNDK 321
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
QY 366 YPSNGCLYSSGWFEDACLSANLNGKYYH-----QKYRGYRNGIFWGTW 410
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 322 F-EGNCAEODGSGWMMNCKHAGHLNGVYGGTYSKASTPNGYDNGIILWATW 372
: : ||||| : : ||||| : : ||||| : : ||||| : : |||||

RESULT 13
US-09-912-741B-1
: Sequence 1, Application US/09912741B
: Patent No. US20020131970A1
: GENERAL INFORMATION:
: APPLICANT: Altieri, Dario C
: APPLICANT: Languino, Lucia R
: APPLICANT: Thornton, George B
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING
: TITLE OF INVENTION: ENDOTHELIAL CELL AND FIBRINOGEN MEDIATED INFLAMMATION
: FILE REFERENCE: 300.1DIV4
: CURRENT APPLICATION NUMBER: US/09/912,741B
: CURRENT FILING DATE: 2001-07-24
: PRIOR APPLICATION NUMBER: US 09/347,877
: PRIOR FILING DATE: 1999-07-06
: PRIOR APPLICATION NUMBER: US 08/748,150
: PRIOR FILING DATE: 1996-11-12
: PRIOR APPLICATION NUMBER: US 08/232,532
: PRIOR FILING DATE: 1994-04-25
: PRIOR APPLICATION NUMBER: US 08/139,562
: PRIOR FILING DATE: 1993-10-19
: PRIOR APPLICATION NUMBER: US 07/898,117
: PRIOR FILING DATE: 1992-06-12
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 411
: TYPE: PRT
: ORGANISM: Unknown
: FEATURE:
: OTHER INFORMATION: expressed
: FEATURE:
: NAME/KEY: CARBOHYD
: LOCATION: 88
: OTHER INFORMATION: site of glycosylation
: FEATURE:
: NAME/KEY: DISULFID
: LOCATION: (153)...(182)
: OTHER INFORMATION: disulfide-bond
: FEATURE:
: NAME/KEY: DISULFID
: LOCATION: (326)...(339)
: OTHER INFORMATION: disulfide-bond
US-09-912-741B-1

Query Match 19.7%; Score 468; DB 10; Length 411;
Best Local Similarity 30.6%; Pred. No. 1.3e-30;
Matches 126; Conservative 54; Mismatches 168; Indels 64; Gaps 13;
```

Db 105 SSIRYQSE-----IYNSNN-----QKIVNLKEKVAQLAQCE-PCKDITVQ 144  
QY 202 SRPVOHLIYKDCSDYAIAGKSSSTYRVTPDPKNSSEYFVCDMTGGMGWTVLQARLDGS 261  
Db 145 ---IHDITGDCODIANKAKOSGLYFIKPLKANQOFLVYCEIDSGNGWIVFQRLDGS 201  
QY 262 TMTFTWQDKAGFGLN-----RRFELGNDKIHLLTKSKEM--ILRIDLEDFNGVELXAL 315  
Db 202 VDFKKNIQYKEGFGHLSLSTQSAIPYALRVELEDNNGRTSTAD 261  
QY 316 YDQFYVANEFLYRLHGVNY-NGTAGDAL--RFNKHYNHDLKFFT-----TPKDNDR 365  
Db 262 YAMFKVGPEDAKYRLTYAYFAGGADGAFDGFDFSDKFFTSHGNGQFSTWONDNDK 321  
QY 366 YFSGNGLYSSGWFDACLSANLNGKYH-----QKRGVRNGIFWGTW 410  
Db 322 F-EGNCAFQDQSGWNNKCHAGLNGVYQGGTYSKASTPNGYDNGIIMATW 372

RESULT 14  
-09-992-598-314  
Sequence 314, Application US/09992598  
Patent No. US20020160384A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Grittisen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC20  
CURRENT APPLICATION NUMBER: US/09/992,598  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07

; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09  
Query Match 19.4%; Score 460.5; DB 9; Length 461;  
Best Local Similarity 28.8%; Pred. No. 6.4e-30;  
Matches 137; Conservative 60; Mismatches 182; Indels 97; Gaps 14;  
QY 7 YWSSAVLATYGFLLVA--NNETEIKDERAKDVCPVLESRGKC-----EEAGECP 56  
Db 30 YVCTVLLALAVLLAVAGVLFVFNHAPAGTAPPVFWSTGAASANSALVTVRADSSH 89  
QY 57 YQVSLPLTIQLPKQFSRIEVEFEVONLKEIVNSLKKSCQCKLQADDNGDPCRNGLLL 116  
Db 90 LSLIDPRCPDLTDSFARLE-----SAQASVLOA-----LT 120  
QY 117 PSTGAPCEVGNRVRELESEVKNL-----SSELKNAKEEINVHGRLL-----EKL 161  
Db 121 EHAQAPRLVGDQEOELDTLADQLPRLARASELQTECMGLRKGHTLGGLSALQSEQ 180  
QY 162 NLVN--MNNIENTYVDSKYANLTFVYNSL--DGKCKSCPSQEQIQ-----SR 203  
Db 181 RLQLLESQGHMAHLVNSVSDILDALQDRGLGRPRNKADLQAPARGTRPRCATGSR 240  
QY 204 PVQHLIYKDCSDYTAIGKRSSEYRYVTPDKNSSFEVYCDMETMGWTVLQARDGSTN 263  
Db 241 P-----RDCLDVLSGGQDDGVYSVFPTHYPAGFYQYCDMRTDGGWTVFORREDGSVN 294  
QY 264 FTRTWQYKAGFNLRRFELWGNDKIHLITKSKEMIILRIDLEDFNGVELYALYDOFYVA - 322  
Db 295 FFRGWDAYRDGFRJTGHEHLGLKRIHALTTQAAAYELHVDLEDFENGATAYARYSGFVGL 354  
QY 323 ---NEFLKYLHYGVNYNGTAGDALRFKNHYNHDLKFEFTTPDKDNDRYPSGNCGLYYSSG 378  
Db 355 FSVDPEDGYPLTVADYSGTAGDSL-----LKHSGMRFTTKDRDSD--HSENNCAAFYGA 408  
QY 379 WPFDACLSANLNGKYYHQYRGVIRMGIFGTWPGVSEAHPPGKYSFKEAKMIRP 434  
Db 409 WMYRNCHTSLNGQYLRGAHASVADGVIEWSWT-----GWQYSLKFSBMLRP 456

## RESULT 15

US-09-989-722-314  
; Sequence 314, Application US/09989722  
; Patent No. US20020072067A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C63  
; CURRENT APPLICATION NUMBER: US/09/989,722

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
8

```

; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      19.4%  Score 460.5; DB 10; Length 461;
Best Local Similarity 28.8%  Pred. No. 6.4e-30;
Matches 137; Conservative 60; Mismatches 182; Indels 97; Gaps 14;

QY 7 YLSSAVLATYGFVVA-NNETEEIKDERAKDVCVRLSRGK-----EEAGECP 56
DB 30 YVLTLLAVLAVLAVTAVGVFLNHAHAPCTAPPPVSTGAASANSALVTVERADSS 89
QY 57 YQVSLPPLTIQPKQFRIEVEFQVONLKEIVNSLKKSCODCKLOADDNGDGRNGLLL 116
DB 90 LSLIDPRCPDLTDSFARLE-----SAQASVLQA-----LT 120
QY 117 PSTGAPGEVGMNRVRELESEVNKL-----SSELKNAKEEINVLHCRLE-----EKL 161
DB 121 EHONQPRLVGDQEQELDTLADQLPRLARASELQTECMGLKRGHGTGQGLSALQSE 180
QY 162 NLVN-MNNTENYVDSKANLTFVNSL--DGKCKPSPQIQ-----SR 203
DB 181 RLQILLESQGHMAHLVNSVSDILDALQDRGLGRPNKADLQAPARGTRPRCATGSR 240
QY 204 PVQHLYIKDCDYVAIGKRSSETYRVTDPKNSSEFVYCDMTGSGWTVLQARLDCSTN 263
DB 241 P-----RCLDLVLLSQODDGVYVFTHYPAGFYQYCDMTDGGWTVFORREDSVN 294
QY 264 FTRTWQDYKAGFNLREFWGLGNDKIHLLTKSKEMILRIDLEDFNGVELYALYDOFYVA- 322
DB 295 FFRGWDAYRDGFGRLTGEHWLGLKRIHALTTQAAAYELHVDLEDFENGTAAYARYGSEFVGL 354
QY 323 ----NEFLKRYLHVNGYNTAGDALREKNHYNHDLKFFTTDPKNDNDYPSGCLYSSG 378
DB 355 FSVDPEDCYPLTVADYSGTAGDSL-----LKHSGMRFTTKDRSD-HSENNCAFYRGA 408
QY 379 WNFDACLSANLNGYHYHQYKTVGRVNGIFGWTWPCVSEAHPGYKSKFEAKMMIRP 434
DB 409 WWTYRNCHT"SNLNGOYLGAHASYADGVEWSSWT-----GWQYSLKFESEMKIRP 456

```

Search completed: December 2, 2002, 07:00:06  
Job time: 14.5374 secs

GenCore version 5.1.3  
Copyright (C) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:49:20 ; Search time 26.1079 Seconds  
(without alignments)  
494.741 Million cell updates/sec

Title: US-09-902-563-2

Perfect score: 2378

Sequence: 1 MGLANNWYLSAVLATYGL.....GYKSFREAKMIRPRHFKP 439

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID                | Description        |
|------------|--------|-------------|--------|-------------------|--------------------|
| 1          | 2378   | 100.0       | 439    | US-09-442-143A-2  | Sequence 2, Appli  |
| 2          | 1853.5 | 77.9        | 432    | US-09-442-143A-4  | Sequence 4, Appli  |
| 3          | 524    | 22.0        | 314    | US-08-525-505A-2  | Sequence 2, Appli  |
| 4          | 524    | 22.0        | 496    | US-09-202-491-7   | Sequence 7, Appli  |
| 5          | 522    | 22.0        | 480    | US-08-740-223A-8  | Sequence 8, Appli  |
| 6          | 522    | 22.0        | 480    | US-09-709-188-8   | Sequence 8, Appli  |
| 7          | 522    | 22.0        | 496    | US-08-373-579-6   | Sequence 6, Appli  |
| 8          | 522    | 22.0        | 496    | US-08-418-595-6   | Sequence 6, Appli  |
| 9          | 522    | 22.0        | 496    | US-08-665-926-6   | Sequence 6, Appli  |
| 10         | 522    | 22.0        | 496    | US-09-162-437-6   | Sequence 6, Appli  |
| 11         | 522    | 22.0        | 496    | US-08-740-223A-6  | Sequence 6, Appli  |
| 12         | 522    | 22.0        | 496    | US-09-351-457-4   | Sequence 4, Appli  |
| 13         | 522    | 22.0        | 496    | US-09-561-500-4   | Sequence 4, Appli  |
| 14         | 522    | 22.0        | 496    | US-09-561-108-4   | Sequence 4, Appli  |
| 15         | 522    | 22.0        | 496    | US-09-351-343-4   | Sequence 4, Appli  |
| 16         | 522    | 22.0        | 496    | US-09-561-526-4   | Sequence 4, Appli  |
| 17         | 522    | 22.0        | 496    | US-09-202-491-8   | Sequence 8, Appli  |
| 18         | 522    | 22.0        | 496    | US-08-817-318-6   | Sequence 6, Appli  |
| 19         | 522    | 22.0        | 496    | US-09-709-188-6   | Sequence 6, Appli  |
| 20         | 521.5  | 21.9        | 498    | US-08-740-223A-20 | Sequence 20, Appli |
| 21         | 521.5  | 21.9        | 498    | US-09-709-188-20  | Sequence 20, Appli |
| 22         | 517.5  | 21.8        | 499    | US-08-740-223A-24 | Sequence 24, Appli |
| 23         | 517.5  | 21.8        | 499    | US-09-709-188-24  | Sequence 24, Appli |
| 24         | 510    | 21.4        | 496    | US-08-740-223A-15 | Sequence 15, Appli |
| 25         | 510    | 21.4        | 496    | US-09-709-188-15  | Sequence 15, Appli |
| 26         | 506.5  | 21.3        | 286    | US-08-960-507-20  | Sequence 20, Appli |
| 27         | 506.5  | 21.3        | 286    | US-09-136-801-20  | Sequence 20, Appli |

```

28 506 21.3 312 1 US-08-525-505A-4 Sequence 4, Appli
29 505 21.2 491 3 US-08-933-821-4 Sequence 4, Appli
30 505 21.2 491 3 US-08-960-507-4 Sequence 4, Appli
31 505 21.2 491 4 US-09-136-828-4 Sequence 4, Appli
32 505 21.2 491 4 US-09-332-928A-4 Sequence 4, Appli
33 505 21.2 491 4 US-09-136-801-4 Sequence 4, Appli
34 505 21.2 491 4 US-09-332-929-4 Sequence 4, Appli
35 498 20.9 496 4 US-08-740-223A-16 Sequence 16, Appli
36 498 20.9 496 4 US-09-709-188-16 Sequence 16, Appli
37 489 20.6 346 3 US-08-960-507-19 Sequence 19, Appli
38 489 20.6 346 4 US-09-136-801-19 Sequence 19, Appli
39 488 20.5 497 4 US-08-740-223A-4 Sequence 4, Appli
40 488 20.5 497 4 US-09-709-188-4 Sequence 4, Appli
41 487 20.5 490 4 US-08-740-223A-12 Sequence 12, Appli
42 487 20.5 496 4 US-09-709-188-12 Sequence 12, Appli
43 487 20.5 496 4 US-08-740-223A-22 Sequence 22, Appli
44 487 20.5 496 4 US-09-709-188-22 Sequence 22, Appli
45 487 20.5 497 1 US-08-373-579-4 Sequence 4, Appli

```

## ALIGNMENTS

```

RESULT 1
US-09-442-143A-2
; Sequence 2, Application US/09442143A
; Patent No. 6403089
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-14
; CURRENT APPLICATION NUMBER: US/09/442.143A
; CURRENT FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens fgl2
US-09-442-143A-2

```

```

Query Match 100.0%; Score 2378; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 5.8e-211;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLANNWYLSAVLATYGLFVYVANNTEETIKDERAKDVCVRLESRGKCEBAGECPYQVS 60
DB 1 MGLANNWYLSAVLATYGLFVYVANNTEETIKDERAKDVCVRLESRGKCEBAGECPYQVS 60
QY 61 LPPLTIQLPKQFSRIEVEFKEVONLKEIVNSLKKSCQCKLQADDNGDGRNGLLPSTG 120
DB 61 LPPLTIQLPKQFSRIEVEFKEVONLKEIVNSLKKSCQCKLQADDNGDGRNGLLPSTG 120
QY 121 APGEVDNRVRELESEVKNLSSELKNKAKKEEINVHGLREKLNLVNMMNIENYVDSKVANL 180
DB 121 APGEVDNRVRELESEVKNLSSELKNKAKKEEINVHGLREKLNLVNMMNIENYVDSKVANL 180
QY 181 TFVNSLDGKCKSCPSQEQTSQRPVQHLYIKDCSDYYAIGKRSSETYRTVTPDKNSSEFV 240
DB 181 TFVNSLDGKCKSCPSQEQTSQRPVQHLYIKDCSDYYAIGKRSSETYRTVTPDKNSSEFV 240
QY 241 YCDMETMGSGTIVLQARLDGSTNFTRTWQDYKAGFNLRRFVWGNLTKHLLTKSEML 300
DB 241 YCDMETMGSGTIVLQARLDGSTNFTRTWQDYKAGFNLRRFVWGNLTKHLLTKSEML 300
QY 301 RIDLEDFNVELYALYDQFYVANEFLKYLRLHWGNYNGTAGDALRFKNKHYNHDLKFFTPD 360
DB 301 RIDLEDFNVELYALYDQFYVANEFLKYLRLHWGNYNGTAGDALRFKNKHYNHDLKFFTPD 360

```

QY 361 KNDYPSGNCGLIYSSGWFDAACLSANLNGKYHOKYRGVRNGIFWGTWPGVSEAHFGG 420  
Db 361 KNDYPSGNCGLIYSSGWFDAACLSANLNGKYHOKYRGVRNGIFWGTWPGVSEAHFGG 420  
QY 421 YKSSFKKAKMIRPKHKP 439  
Db 421 YKSSFKKAKMIRPKHKP 439

## RESULT 2

US-09-442-143A-4  
; Sequence 4, Application US/09442143A  
; Patent No. 6403089  
; GENERAL INFORMATION:  
; APPLICANT: Levy, Gary  
; APPLICANT: Clark, David A.  
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation  
; FILE REFERENCE: 9579-14  
; CURRENT APPLICATION NUMBER: US/09/442,143A  
; PRIOR FILING DATE: 1999-11-15  
; PRIOR APPLICATION NUMBER: US 60/046,537  
; PRIOR FILING DATE: 1997-05-17  
; PRIOR APPLICATION NUMBER: US 60/061,684  
; PRIOR FILING DATE: 1997-10-10  
; NUMBER OF SEQ. ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ. ID NO. 4  
; LENGTH: 432  
; TYPE: PRT  
; ORGANISM: Murine fg12  
US-09-442-143A-4

Query Match 77.9%; Score 1853.5; DB 4; Length 432;  
Best Local Similarity 77.7%; Pred. No. 1.3e-162;  
Matches 341; Conservative 42; Mismatches 49; Indels 7; Gaps 4;  
QY 1 MCLANWYLSAVLATYGLVANNETEEIKDKRAKDCVPRLESRGKCEAGECPYQVS 60  
Db 1 MCLANWYLSAVLATYGLVANNETEEIKDKRAKDCVPRLESRGKCEAGECPYQVS 60  
QY 61 LPTLTIQPLQSGMEVLKVRYLKAVDSLKSCQCKLQADDDHDPGNG----GNG 114  
Db 59 LPTLTIQPLQSGMEVLKVRYLKAVDSLKSCQCKLQADDDHDPGNG----GNG 114  
QY 121 APGEVGNVRLESEVNKLSELNAKEEINVLHGRLEKLNLMNMNIENYVDSKVANL 180  
Db 115 AP-TAEDSRVQELSESQVKNLSSELNAKEDQIQGLQRLTLHLVNMNIENYVDSKVANL 173  
QY 181 TFVNSLDGKCKSCPSQEQISRPVQHLIYKDCSDYVYAGKRSSTYRTVTPDKNSSFEV 240  
Db 174 TFVNSLDGKCKSCPSQEQISRPVQHLIYKDCSDYVYAGKRSSTYRTVTPDKNSSFEV 233  
QY 241 YCDMETGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRFVLGNDKIHLLTKSKEMIL 300  
Db 234 YCDMETGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRFVLGNDKIHLLTKSKEMIL 293  
QY 301 RDLDELNGVLYALYDQFYVANEFLKYLHVGNYNGTAGDALREKHYNHDLKFTTPD 360  
Db 294 RDLDELNGVLYALYDQFYVANEFLKYLHVGNYNGTAGDALREKHYNHDLKFTTPD 353  
QY 361 KNDYPSGNCGLIYSSGWFDAACLSANLNGKYHOKYRGVRNGIFWGTWPGVSEAHFGG 420  
Db 354 KNDYPSGNCGLIYSSGWFDAACLSANLNGKYHOKYRGVRNGIFWGTWPGVSEAHFGG 413  
QY 421 YKSSFKKAKMIRPKHKP 439  
Db 414 YKSSFKKAKMIRPKHKP 432

## RESULT 3

US-08-525-505A-2  
; Sequence 2, Application US/08525505A

; Patent No. 5807711  
; GENERAL INFORMATION:  
; APPLICANT: HARA, HIROSHI  
; APPLICANT: YOSHIMURA, HIROMITSU  
; APPLICANT: MATSUKI, YUMIKO  
; APPLICANT: SHINDO, SAEKO  
; APPLICANT: HANADA, KAZUNORI  
; TITLE OF INVENTION: PARENCHYMAL HEPATOCYTE GROWTH SUBSTANCE  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/525,505A  
; FILING DATE: 22-SEP-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP94/00455  
; FILING DATE: 22-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 05-063905  
; FILING DATE: 23-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 4587-012-0 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 314 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-525-505A-2

Query Match 22.0%; Score 524; DB 1; Length 314;  
Best Local Similarity 37.6%; Pred. No. 3e-40;  
Matches 124; Conservative 46; Mismatches 104; Indels 56; Gaps 13;  
QY 125 VGD-NRVRE---LESEVNKLSELNAKEEINVLHGRLEKLNLMNMNIENYVDSKVAN 179  
Db 23 LGDENCLQEQVRLRAQVRLQTRVKKQQVVIQAQLLHEK--EVQFLDRGQEDSFID---- 75  
QY 180 LTFVNSLDGKCKSCPSQEQISRPVQHLIYKDCSDYVYAGKRSSTYRTVTPDKNSSFE 239  
Db 76 -----LGK-----RH--YADCSEIYNDGFKHSGFYKIKPLQSLAEFS 111  
QY 240 VYCDMETGGGWTVLQARLDGSTNFTRTWQDYKAGFNL---RREFWLNDKIHLLTKSK 296  
Db 112 VYCDMSD-GGGTVTIQRSDSGSENFNRGNDYENGFGNFVQSGNCEYWLGNKINLLTMOG 170  
QY 297 EMILRIDLEDNFGVLYALYDQFYVANEFLKYLHVGNYNGTAGDALREKHYNHDLKFTTPD 350  
Db 171 DYTLKIDLTFDEKNSRFAQYKEFKVGDEKSFYELNTEYSGTAGDSLSCTPHPEVQHWAS 230  
QY 351 HDLKFPTTDPKNDYPSGNCGLIYSSGWFDAACLSANLNGKYHOKYRG-VRNGIFWGT 409  
Db 231 HOTMKFSTRDRNDNY-NGNCAEEQSGWGFNRCHSANLNGVYQGYRAETDNGVYWT 289  
QY 410 WPGVSEAHFGGKYKSKFEAKMIRPKHKP 439

Db 290 W-----RGWYSLKSVVMKIRPSDEIP 311

## RESULT 4

US-09-202-491-7  
; Sequence 7, Application US/09202491  
; Patent No. 6432667  
; GENERAL INFORMATION:  
; APPLICANT: Valenzuela et al.  
; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF  
; FILE REFERENCE: REG330-K  
; CURRENT APPLICATION NUMBER: US/09/202,491  
; CURRENT FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: PCT/US97/10728  
; EARLIER FILING DATE: 1997-06-19  
; EARLIER APPLICATION NUMBER: 60/022,999  
; EARLIER FILING DATE: 1996-08-02  
; EARLIER APPLICATION NUMBER: 60/021,087  
; EARLIER FILING DATE: 1996-07-02  
; EARLIER APPLICATION NUMBER: 08/665,926  
; EARLIER FILING DATE: 1996-06-19  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-202-491-7

Query Match 22.0%; Score 524; DB 4; Length 496;  
Best Local Similarity 32.9%; Pred. No. 6e-40;  
Matches 127; Conservative 65; Mismatches 154; Indels 40; Gaps 9;

QY 61 LPPLTIQPKQFSRIEVEFKEVQNLKELVNSLKKSCQDCKLQADDNGDPGRNGLLPSTG 120  
Db 142 LTDVEAQLNQTTRL-----ELQLQHSISTNKLEKQILDQTSINKLHNKSF----- 191  
QY 121 APOGEVDNRVELESEVKNLSSELNAKEEINVLHGRL-----EKLNLVNMNNIENY 172  
Db 192 -----EOKVLDMGKHSEQLQSMKEQKDELQVLVSKQSSVDELEKLVATVNN--SL 243  
QY 173 VDSKVANITFVNSLDGCKSPCEQIQSRPVQHLIYKDCSYAIGKRSSTYRTPTD 232  
Db 244 LKQOQHDLMETVNSLLTMSSPNSKSSVAIRKEQITFDRAEIPKSLTSGIYLTTP 303  
QY 233 PKNSFEVYCDMETGGGWTVLQALDGSNTFTWQDYKAGFNLRFELWLNKIHLL 292  
Db 304 NSTEIKAYCDMDVGGGWTVIQHREDGSDVDFQWTKEYKEGFGNPLGEVWLNFEVSQL 363  
293 TKSKEMLIRIDLEDFNGVELYALYQFYVANEFLYRLHVGNYNGTAGDALRFNKHYND 352  
Db 364 TGQHYRLKIQLDWEGNEAHSILYHFLAGEESNYRIHLTGLTGAAKISSISQPGSD- 422  
QY 353 LKFFTPDKDNDYPSNGCLYSSGWFDFACLSANLNGKYYHQYRGVR-NGIFWGTWP 411  
Db 423 ---FSTKQSDNDKIC-KCSQMLSGGWFDFACGPNLNGQYYPQKQNTNFNGIKRYWK 478  
QY 412 GVSEAHPGGYSKSPFKAAMKPKHF 437  
Db 479 G-----SGY--SLKATMTMIRPADF 496

## RESULT 5

US-08-740-223A-8  
; Sequence 8, Application US/08740223A  
; Patent No. 6265564  
; GENERAL INFORMATION:  
; APPLICANT: Davis, et al.  
; TITLE OF INVENTION: Expressed Ligand - Vascular  
; TITLE OF INVENTION: Intercellular Signalling Molecule  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.

STREET: 777 Old Saw Mill Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/740,223A  
FILING DATE: 25-OCT-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION NUMBER: USSN 60/022/999  
FILING DATE: 02-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Cobert, Robert J  
REGISTRATION NUMBER: 36,108  
REFERENCE/DOCKET NUMBER: REG 333  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 480 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Mature TL2 protein  
LOCATION: 1..480  
OTHER INFORMATION:  
US-08-740-223A-8

Query Match 22.0%; Score 522; DB 4; Length 480;  
Best Local Similarity 27.9%; Pred. No. 8.7e-40;  
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRKCE--EAGECPYQVSLPPL-----TIQ--LPKQFSRIEVEFKEVON 84  
Db 10 MDSICKKQYQVHGSSCYTFLPEMDNCRSSSPYVSNVADAPLETDDSVQRLQVLEN 69  
QY 85 LKE-----IVNSLKKSCQDCKLQADDNGDPGRNGLLPSTGAPCEVGDN--- 128  
Db 70 IMENNTQMLKLENYIODNNKKEVETQONAVON-----QTAVMEIGTLLNQ 118  
QY 129 -----RVRELESEV-----NKLSSSELKNAAKEEINVLHGR---LEKL 161  
Db 119 TAEQTRKLTDEAQLVQNLQTRLEQLLEHLSLSTNLEKQILDQTSINKLODKNSFLEKK 178  
QY 162 NLV-----NMNNIENYVDSKVANLTF-----V 183  
Db 179 VLAMEDKHIIQLOSIKEEKDQLQVLVSKQNSIIELEKKIVTATVNNSVLQKQOHDLMET 238  
QY 184 VNSLDGCKSPCEQIQSRPVQHLIYKDCSYAIGKRSSTYRTPTDPKNSFEVCD 243  
Db 239 VNNLLTMSTNSAKDPTVAKEEQISFDCAEYFKSGHTTNGIYTLTPNSTEIRKAYCD 298  
QY 244 METGGGWTVLQALDGSNTFTWQDYKAGFNLRFELWLNKIHLLTKSKEMILRID 303  
Db 299 MEAGGGWTIIQRRREDGSDVDFQWTKEYKVGFGNPSGEYWLNGNEFVSQLTNOQRYVLKIH 358  
QY 304 LEDFNGVELYALYQFYVANEFLYRLHVGNYNGTAGDALRFNKHYNDLKFETTPDKDN 363  
Db 359 LKDWEGNEAHSYEHFYLSEELNRYHLKGLGTAGKISSISQPGND-----FSTKQSDN 414  
QY 364 DRYPSGNGCLYSSGWFDFACLSANLNGKYYHQYRGVR-NGIFWGTWPGVSEAHFPGYK 422  
Db 415 DKCIC-KCSQMLTGGWDFDAGCPSNLNGMYYPQKQNTNFNGIKRYWKG-----SGY- 466

QY 423 SSFKEAKMIRPKHF 437  
| | | | |  
Db 467 -SLKATTMMIRPADF 480

## RESULT 6

US-09-709-188-8  
; Sequence 8, Application US/09709188  
; Patent No. 6441137

## GENERAL INFORMATION:

; APPLICANT: Davis et al.  
; TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling Molecule  
; FILE REFERENCE: REG 333-2  
; CURRENT APPLICATION NUMBER: US/09/709,188  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 08/740,223  
; PRIOR FILING DATE: 1996-10-25  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-709-188-8

Query Match 22.0%; Score 522; DB 4; Length 480;

Best Local Similarity 27.9%; Pred. No. 8,7e-40;

Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGGCPQVSLPPL-----TIQ--LPKQFSRIEEVPKEVQN 84

Db 10 MDSIGKKQYQVHGSCSYTLLPEMNCRSSSPVYVNAVQRPADLEYDDSVQRLQVLEN 69

QY 85 LKE-----IVNSLKKSCDCKLQADDNGDPGRNGLLLPSTGAPGEVGN---- 128

Db 70 IMENNTQWLKLENYIQDNMKKEMVEIQONAVON-----QTAVWIEIGTNLLNQ 118

QY 129 -----RVRELESEV-----NKLSELKNAKEEINVLHGR---LEKL 161

Db 119 TAEQTRKLTDEVAQVNLQTTRELEQLLEHSLSTNKLKQILDQTSINKLQDKNSFLEKK 178

QY 162 NLV-----NMNNTENYVDSKVANLTF-----V 183

Db 179 VLAMEDKHIIQLQSIKEEKDQQLVLSVKQNSIIIELEKKIVTATVNNVSLQKQHDLMET 238

QY 184 VNSLDGKCKSCPSQEQIOSRPVQHLYIKDCSDYIAGKRSSEYRYVTPDPKNSFEVYCD 243

Db 239 VNNLLTMMSTNSAKDPTVAKEEQISFRDCAEVKFSGHTTNGIYTLTFPNSTEEIKAYCD 298

QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRREFWLGNDKIHLLTKSKEMILRID 303

Db 299 MEAGGGWTIIQRREDGSDVDFQRTWKYKVGFGNPSGEYWLGNFYSQLTNQORYVLKIH 358

QY 304 LEDPNGVELYALDYQVANEFLYRLHVCNVTNGTAGDALRPNKHYNHDLKFFPTDPKON 363

Db 359 LKDWEGNEAYSLYEHFYLSSSELNYRIHLKGLTGKTAGKISSISQPGND----FSTKGDGN 414

QY 364 DRYPSGCGGLYSSGWFDACLSANLNGKYTHQYKRGVR-NGIFWGTWPGVSEAHPPGYK 422

Db 415 DKCIC-KCSQMLTGGWFDACGPNLNGYTFQRTNKNKNGIKWYWKG-----SGY- 466

QY 423 SSFKEAKMIRPKHF 437

Db 467 -SLKATTMMIRPADF 480

## RESULT 7

US-08-373-579-6

; Sequence 6, Application US/08373579

; Patent No. 5650490

## GENERAL INFORMATION:

; APPLICANT: Davis, et al.  
; TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Regeneron Pharmaceuticals, Inc.

; STREET: 777 Old Saw Mill River Road

; CITY: Tarrytown

; STATE: New York

; COUNTRY: USA

; ZIP: 10591

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/373,579

; FILING DATE: 17-JAN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/353,503

; FILING DATE: 09-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/348,492

; FILING DATE: 02-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/330,261

; FILING DATE: 27-OCT-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/319,932

; FILING DATE: 07-OCT-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Cobert, Robert J.

; REGISTRATION NUMBER: 36,108

; REFERENCE/DOCKET NUMBER: REG 330-D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (914) 345-7400

; TELEFAX: (914) 345-7721

; INFORMATION FOR SEQ ID NO:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 496 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-373-579-6

Query Match 22.0%; Score 522; DB 1; Length 496;

Best Local Similarity 27.9%; Pred. No. 9,1e-40;

Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGGCPQVSLPPL-----TIQ--LPKQFSRIEEVPKEVQN 84

Db 26 MDSIGKKQYQVHGSCSYTLLPEMNCRSSSPVYVNAVQRPADLEYDDSVQRLQVLEN 85

QY 85 LKE-----IVNSLKKSCDCKLQADDNGDPGRNGLLLPSTGAPGEVGN---- 128

Db 86 IMENNTQWLKLENYIQDNMKKEMVEIQONAVON-----QTAVWIEIGTNLLNQ 134

QY 129 -----RVRELESEV-----NKLSELKNAKEEINVLHGR---LEKL 161

Db 135 TAEQTRKLTDEVAQVNLQTTRELEQLLEHSLSTNKLKQILDQTSINKLQDKNSFLEKK 194

QY 162 NLV-----NMNNTENYVDSKVANLTF-----V 183

Db 195 VLAMEDKHIIQLQSIKEEKDQQLVLSVKQNSIIIELEKKIVTATVNNVSLQKQHDLMET 254

QY 184 VNSLDGKCKSCPSQEQIOSRPVQHLYIKDCSDYIAGKRSSEYRYVTPDPKNSFEVYCD 243

Db 255 VNNLLTMMSTNSAKDPTVAKEEQISFRDCAEVKFSGHTTNGIYTLTFPNSTEEIKAYCD 314

QY 244 METMGGGWTVLQARLDGSTNFTRTWQDYKAGFNLRRREFWLGNDKIHLLTKSKEMILRID 303

Db 315 MEAGGGWTIIQRREDGSDVDFQRTWKYKVGFGNPSGEYWLGNFYSQLTNQORYVLKIH 374



US-08-665-926-6

Query Match 22.0%, Score 522; DB 2; Length 496;  
Best Local Similarity 27.9%; Pred. No. 9.1e-40;  
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRKGCE--EAGECPQVSLPPL-----TIQ--LPKQFSRIEEVKEVQN 84  
Db 26 MDSIGKKQYQVHGSCSTFLPEMNCRSSSSPYVNAVORDAPLEYDDSVQRLQVLEN 85  
QY 85 LKE-----IVNSLKKSCQCKLOADDNGDPRNGLLLPSTGAPGEVGN---- 128  
Db 86 IMENNTOWLMKLENYIQDNMKEMVEIQNAVN-----QTAVMIEIGTNLLNQ 134  
QY 129 -----RVRELESEV-----NKLSELKNAKEEINVLHGR---LEKL 161  
Db 135 TAEQTRKLTDEAQLVQVLTOTTRLEQLLLEHSLSTNKLKQILDQISEINKLQDNKSFLEKK 194  
QY 162 NLV-----NMNNTIENYVDSKVANLTF-----V 183  
Db 195 VLAMEDKHIIQLOSIKEEKDQQLVLSKQNSIIIELEKIVTATVNNSVLQKQHDLMET 254  
QY 184 VNSLDGCKSCPSQEQIOSRPVQHLIYKDCSDYYAIGKRSSEYRVTPDPKNSSEVYCD 243  
Db 255 VNNLLTMNSTNSAKDPTVAKEQISFRDCAEVKSGHTTNGIYTLTFPNSTEEIKAYCD 314  
QY 244 METMGGGWTIQLARLDGSTNFTTWQDYKAGFGLNRREFWGLNDKIHLLTKSKEMILRID 303  
Db 315 MEAGGGWTIIQRREDGSDVDFQRTWKEYKVGFGNPSGEYWLGNFVSQLTNQORYVLKIH 374  
QY 304 LEDFNGVELYALDOFYVANEFLKYLHVGNVNTAGDALRPNKHYNHDLKFFTTDPKDN 363  
Db 375 LKDWEGNEAYSLYEHYLSSEELNRIHLKGLTGTAGKISSISQPCND----FSTKQDGN 430  
QY 364 DRYPSGNGGLYSSGWWFDCALSANLNGKYHYQKYGVR-NGIFWGTWPGVSEAHPGYK 422  
Db 431 DKCIC-KCSQMLTGGWFDACGPNLNGMYYPQONTNKGKIKWYKRG-----SGY- 482  
QY 423 SSFKEAKMIRPKHF 437  
Db 483 -SLKATTMIRPADF 496

Query Match 22.0%, Score 522; DB 4; Length 496;  
Best Local Similarity 27.9%; Pred. No. 9.1e-40;  
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRKGCE--EAGECPQVSLPPL-----TIQ--LPKQFSRIEEVKEVQN 84  
Db 26 MDSIGKKQYQVHGSCSTFLPEMNCRSSSSPYVNAVORDAPLEYDDSVQRLQVLEN 85  
QY 85 LKE-----IVNSLKKSCQCKLOADDNGDPRNGLLLPSTGAPGEVGN---- 128  
Db 86 IMENNTOWLMKLENYIQDNMKEMVEIQNAVN-----QTAVMIEIGTNLLNQ 134  
QY 129 -----RVRELESEV-----NKLSELKNAKEEINVLHGR---LEKL 161  
Db 135 TAEQTRKLTDEAQLVQVLTOTTRLEQLLLEHSLSTNKLKQILDQISEINKLQDNKSFLEKK 194  
QY 162 NLV-----NMNNTIENYVDSKVANLTF-----V 183  
Db 195 VLAMEDKHIIQLOSIKEEKDQQLVLSKQNSIIIELEKIVTATVNNSVLQKQHDLMET 254  
QY 184 VNSLDGCKSCPSQEQIOSRPVQHLIYKDCSDYYAIGKRSSEYRVTPDPKNSSEVYCD 243  
Db 255 VNNLLTMNSTNSAKDPTVAKEQISFRDCAEVKSGHTTNGIYTLTFPNSTEEIKAYCD 314  
QY 244 METMGGGWTIQLARLDGSTNFTTWQDYKAGFGLNRREFWGLNDKIHLLTKSKEMILRID 303  
Db 315 MEAGGGWTIIQRREDGSDVDFQRTWKEYKVGFGNPSGEYWLGNFVSQLTNQORYVLKIH 374  
QY 304 LEDFNGVELYALDOFYVANEFLKYLHVGNVNTAGDALRPNKHYNHDLKFFTTDPKDN 363  
Db 375 LKDWEGNEAYSLYEHYLSSEELNRIHLKGLTGTAGKISSISQPCND----FSTKQDGN 430  
QY 364 DRYPSGNGGLYSSGWWFDCALSANLNGKYHYQKYGVR-NGIFWGTWPGVSEAHPGYK 422  
Db 431 DKCIC-KCSQMLTGGWFDACGPNLNGMYYPQONTNKGKIKWYKRG-----SGY- 482  
QY 423 SSFKEAKMIRPKHF 437  
Db 483 -SLKATTMIRPADF 496

RESULT 11  
US-08-740-223A-6  
; Sequence 6, Application US/08740223A  
; Patent No. 6265564  
; GENERAL INFORMATION:

US-09-162-437-6  
; Sequence 6, Application US/09162437  
; Patent No. 6166185  
; GENERAL INFORMATION:  
; APPLICANT: Davis, et al.  
; TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
; STREET: 777 Old Saw Mill River Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10591  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/162.437  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/418,595  
; FILING DATE: 06-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/373,579



## RESULT 13

US-09-561-500-4  
; Sequence 4, Application US/09561500  
; Patent No. 6342219  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002500  
; CURRENT APPLICATION NUMBER: US/09/561,500  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-561-500-4

Query Match 22.08; Score 522; DB 4; Length 496;  
Best Local Similarity 27.94; Pred. No. 9.1e-40;

Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGCPYQVSLPPL-----TIQ---LPKQFSRIEVEFKVQN 84  
DB 26 MDSIGKKQYQVHGSCSYTELLPEMDCRSSSPVSVNAVORDAPLEYDSDVORLQVLEN 85  
QY 85 LKE-----IVNSLKKSCQCKLOADDNGDPRNGLLLPSTGAPGEVGN---- 128  
DB 86 IMENNTQWLKMLNVIQDNMKKEWVIOQNAVON-----QTAVMTEIGTNLLNQ 134  
QY 129 -----RVRELESEV-----NKLSELKNAKEINVLHCR---LEKL 161  
DB 135 TAEQTRKLTDEVAQVLNQTTRLELQLEHLSLSTNKLKQILDQTSINKLQDNKSFLEKK 194  
QY 162 NLV-----NNNINENYVDSKANLTF-----V 183  
DB 195 VLAMEDKHIIQLOSIKEEKQLOLVSKNSIIELEKKIVTATVNSVLOKQOHDLMET 254  
QY 184 VNSLDKCKSKCPQSQISRPVQHLIYKDCSDYAIAGKRSSEYRYVTPDKNSSFEVYCD 243  
DB 255 VNNLLTMMSTNSAKDPTVAKEQISFRDCAEVKSGHTTNGIYLTLPNSTEIKAYCD 314  
QY 244 METMGCGTIVLQARLDGSTNFTWQDYKAGFNLRFELWLGNDKHLITLTKSKEMILRID 303  
DB 315 MEAGGGWTIIORREDGSDVDFQRTWKYKVGFGNPSGEYWLGNFVSQLTNOQRYVLKIH 374  
QY 304 LEDFNGVELYALDYQYVANEFLKYLHVGNNGTAGDALRNFKNHYNHDLKFFTPDKON 363  
DB 375 LKDWEGNEAYSLEYEHFVLSSEELNYRIHLGLTGTACKISSISQPCND----FSTKDGDN 430  
QY 364 DRYPSGNCGLYSSGWDFDACLANSNGKYYHOKYRCVR-NGIFWGTWPGVSEAHGPGYK 422  
DB 431 DKCIC-KCSQMLTGGWDFDAGGFSNLNGMYYPQONTNFKNGIKWYKKG-----SGY- 482  
QY 423 SSFKEAKMMIRPKHF 437  
DB 483 -SLKATTMMIRPADF 496

## RESULT 14

US-09-561-108-4  
; Sequence 4, Application US/09561108  
; Patent No. 6342221  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002584  
; CURRENT APPLICATION NUMBER: US/09/561,108

; CURRENT FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-561-108-4

Query Match 22.08; Score 522; DB 4; Length 496;  
Best Local Similarity 27.94; Pred. No. 9.1e-40;

Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

QY 43 LESRGKCE---EAGCPYQVSLPPL-----TIQ---LPKQFSRIEVEFKVQN 84  
DB 26 MDSIGKKQYQVHGSCSYTELLPEMDCRSSSPVSVNAVORDAPLEYDSDVORLQVLEN 85  
QY 85 LKE-----IVNSLKKSCQCKLOADDNGDPRNGLLLPSTGAPGEVGN---- 128  
DB 86 IMENNTQWLKMLNVIQDNMKKEWVIOQNAVON-----QTAVMTEIGTNLLNQ 134  
QY 129 -----RVRELESEV-----NKLSELKNAKEINVLHCR---LEKL 161  
DB 135 TAEQTRKLTDEVAQVLNQTTRLELQLEHLSLSTNKLKQILDQTSINKLQDNKSFLEKK 194  
QY 162 NLV-----NNNINENYVDSKANLTF-----V 183  
DB 195 VLAMEDKHIIQLOSIKEEKQLOLVSKNSIIELEKKIVTATVNSVLOKQOHDLMET 254  
QY 184 VNSLDKCKSKCPQSQISRPVQHLIYKDCSDYAIAGKRSSEYRYVTPDKNSSFEVYCD 243  
DB 255 VNNLLTMMSTNSAKDPTVAKEQISFRDCAEVKSGHTTNGIYLTLPNSTEIKAYCD 314  
QY 244 METMGCGTIVLQARLDGSTNFTWQDYKAGFNLRFELWLGNDKHLITLTKSKEMILRID 303  
DB 315 MEAGGGWTIIORREDGSDVDFQRTWKYKVGFGNPSGEYWLGNFVSQLTNOQRYVLKIH 374  
QY 304 LEDFNGVELYALDYQYVANEFLKYLHVGNNGTAGDALRNFKNHYNHDLKFFTPDKON 363  
DB 375 LKDWEGNEAYSLEYEHFVLSSEELNYRIHLGLTGTACKISSISQPCND----FSTKDGDN 430  
QY 364 DRYPSGNCGLYSSGWDFDACLANSNGKYYHOKYRCVR-NGIFWGTWPGVSEAHGPGYK 422  
DB 431 DKCIC-KCSQMLTGGWDFDAGGFSNLNGMYYPQONTNFKNGIKWYKKG-----SGY- 482  
QY 423 SSFKEAKMMIRPKHF 437  
DB 483 -SLKATTMMIRPADF 496

## RESULT 15

US-09-351-543-4  
; Sequence 4, Application US/09351543  
; Patent No. 6406693  
; GENERAL INFORMATION:  
; APPLICANT: THORPE, PHILIP E.  
; APPLICANT: RAN, SOPHIA  
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING ANTIBODIES TO  
; FILE REFERENCE: 4001.002200  
; CURRENT APPLICATION NUMBER: US/09/351,543  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-351-543-4

Query Match 22.08; Score 522; DB 4; Length 496;

```

Best Local Similarity 27.9%; Pred. No. 9.1e-40;
Matches 138; Conservative 74; Mismatches 159; Indels 124; Gaps 15;

Qy 43 LESRGKCE--EAGECPYQVSLPPL-----TIQ--LPKQFSRIEVEFKEVON 84
Db 26 MDSIGKQYQVHGSCSYTFLLPENDNCRSSSPYVSNVQRPDAFLEYDDSVQRLQVLEN 85
Qy 85 LKE-----JVSILKSCDCKLOADDNGDPGRNGLLLPSTGAPGEVGDN---- 128
Db 86 IMENTQWMLKENYIQDNKKKEWVEIQONAVON-----OTAVMIEIGTNLLNQ 134
Qy 129 -----RVRELESFV-----NKLSELKNAAKEEINVILHGR---LEKL 161
Db 135 TAEQTRKLTDEAQLVNOTTRLEQLLEHLSLSTNKLKQILDQISEINKLQDKNSFLEKK 194
Qy 162 NLV-----NMNNIENYVDSKVANLTF-----V 183
Db 195 VLAMEDKHILOSIKEEKDQLOVLVSKQNSIIEELEKKIVTATVNNNSVLQKQOHDLMET 254
Qy 184 VNSLDGKCKSPQEQIQSRPVOHLIYKDCSDYYAIGKRSETYRVTPDPKNSSEFVYCD 243
Db 255 VNNLLTMSTSNSAKDPTVAKEEQISFRDCAEVFKSGHTTNGIYTLTPNASTEIEIKAYCD 314
Qy 244 METMGSGWTVLQARLDGSTNFTRTWODYKAGFGNLRREEWLGNMDKIHLTKSKEMILRD 303
Db 315 MEAGGGWTIIORREDGSDVDFORTWKEKYKVGFGNPSGEYWLGNFEVSQLTNQORYVLIK 374
Qy 304 LEDFNGVELYALYDQFYVAMEFLKYRLHVGNYNGTAGDALRFNKHYNHDLKFFETTPDKDN 363
Db 375 LKDWEGNEAYSIIYEHFYLSEELNYRIHLKGLTGTAGKISSISQPGND---ESTKGDGN 430
Qy 364 DRYPSGNGCLYSSGWNFACLSANLINGKYHOKYRGVR-NGIFWGTWPGVSEAHFPGYK 422
Db 431 DKCIC-KSQMLTGGWFWDCGFSMLNGMYYPQRQNTNFKNGIKWYWKG-----SGY- 482
Qy 423 SSFKEAKWIRPKHF 437
Db 483 -SLKATTWIRPADF 496

```

Search completed: December 2, 2002, 06:54:34  
Job time : 28.1079 secs

1

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:48:57 ; Search time 1.05727 Seconds  
(without alignments)  
1363.907 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRVPSGNCGLYSSG 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134222 residues

total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 88    | 100.0       | 432    | 2 I56934 | fibrinogen-like pr |
| 2          | 88    | 100.0       | 432    | 2 A27447 | cytotoxic T-lympho |
| 3          | 88    | 100.0       | 439    | 2 I37391 | fibrinogen-like pr |
| 4          | 46    | 52.3        | 2403   | 2 A59386 | sanko - human      |
| 5          | 46    | 52.3        | 3461   | 2 S58870 | reelin precursor - |
| 6          | 43    | 48.9        | 225    | 2 C82893 | hypothetical prote |
| 7          | 43    | 48.9        | 284    | 2 I49707 | germ cell specific |
| 8          | 43    | 48.9        | 339    | 2 T24007 | hypothetical prote |
| 9          | 43    | 48.9        | 436    | 2 C86624 | hypothetical prote |
| 10         | 43    | 48.9        | 456    | 2 H72000 | hypothetical prote |
| 11         | 42.5  | 48.3        | 696    | 2 G71829 | probable outer mem |
| 12         | 41    | 46.6        | 273    | 2 T16246 | hypothetical prote |
| 13         | 41    | 46.6        | 368    | 2 A81289 | hypothetical prote |
| 14         | 41    | 46.6        | 385    | 2 T26404 | hypothetical prote |
| 15         | 41    | 46.6        | 437    | 2 A72513 | hypothetical prote |
| 16         | 41    | 46.6        | 517    | 1 ERAD47 | early E2A DNA-bind |
| 17         | 41    | 46.6        | 591    | 1 WBP02  | gene P2 protein -  |
| 18         | 41    | 46.6        | 680    | 2 H70347 | outer membrane pro |
| 19         | 41    | 46.6        | 798    | 2 T00131 | xylan 1,4-beta-xyl |
| 20         | 40    | 45.5        | 92     | 2 A72242 | ferredoxin - Therm |
| 21         | 40    | 45.5        | 100    | 2 JN0734 | hypothetical 11.49 |
| 22         | 40    | 45.5        | 165    | 2 D86894 | hypothetical prote |
| 23         | 40    | 45.5        | 245    | 2 G64210 | uracil DNA glycosy |
| 24         | 40    | 45.5        | 286    | 2 A33546 | actin-capping prot |
| 25         | 40    | 45.5        | 294    | 2 AF3043 | calcium-binding pr |
| 26         | 40    | 45.5        | 305    | 2 F98242 | regucalcin (AB0379 |
| 27         | 40    | 45.5        | 455    | 2 E64454 | proline-trna ligas |
| 28         | 40    | 45.5        | 474    | 1 ERAD41 | early E2A DNA-bind |
| 29         | 40    | 45.5        | 484    | 1 ERAD12 | early E2A DNA-bind |

## ALIGNMENTS

## RESULT 1

I56934

fibrinogen-like protein - mouse

C:Species: Mus sp (mouse)

C:Date: 28-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 03-Dec-1999

C:Accession: I56934

R:Part, R.L.; Fung, L.; Reneker, J.; Myers-Mason, N.; Leibowitz, J.L.; Levy, G.

J. Virol. 69, 5033-5038, 1995

A:Title: Association of mouse fibrinogen-like protein with murine hepatitis virus-ind

A:Reference number: I56934; MUID:95333285; PMID:7609073

A:Accession: I56934

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-432 &lt;RES&gt;

A:Cross-references: GB:S78773; NID:g1042169; PIDN:AA34823.1; PID:g1042170

C:Genetics:

A:Gene: musfibip

C:Superfamily: fibrinogen gamma chain: fibrinogen beta/gamma homology

F:203-428/Domain: fibrinogen beta/gamma homology &lt;FBG&gt;

Query Match 100.0%; Score 88; DB 2; Length 432;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRVPSGNCGLYSSG 15

|||||

Db 357 DRVPSGNCGLYSSG 371

## RESULT 2

A27447

cytotoxic T-lymphocyte-specific protein precursor (clone pT49) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 03-Dec-1999

C:Accession: A27447

R:Koyama, T.; Hall, L.R.; Hasegawa, W.G.; Toneyawa, S.; Saito, H.

Proc. Natl. Acad. Sci. U.S.A. 84, 1609-1613, 1987

A:Title: Structure of a cytotoxic T-lymphocyte-specific gene shows a strong homology

A:Reference number: A27447; MUID:87175527; PMID:3550794

A:Accession: A27447

A:Molecule type: mRNA

A:Residues: 1-432 &lt;KOY&gt;

A:Cross-references: GB:M16238; NID:q193304; PIDN:AAA37624.1; PID:g387156

C:Superfamily: fibrinogen gamma chain: fibrinogen beta/gamma homology

F:203-428/Domain: fibrinogen beta/gamma homology &lt;FBG&gt;

Query Match 100.0%; Score 88; DB 2; Length 432;  
Best Local Similarity 100.0%; Pred. No. 2e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRVPSGNCGLYSSG 15

|||||

Db 357 DRYPGNGGLYSSG 371

## RESULT 3

I37391

fibrinogen-llke protein expressed in T lymphocytes (p749) - human

C:Species: Homo sapiens (man)

C:Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 21-Jul-2000

C:Accession: I37391; S47273

R:Ruegg, C.; Pytela, R.

Gene 160, 257-262, 1995

A:Title: Sequence of a human transcript expressed in T-lymphocytes and encoding a fibrin

A:Reference number: I37391; MUID:95369700; PMID:7642106

A:Accession: I37391

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-439 <RES>

A:Cross-references: EMBL:236531; NID:g535184; PID:CA85298.1; PID:g535185

A:Note: submitted to the EMBL Data Library, August 1994

Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology

210-435/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 100.0%; Score 88; DB 2; Length 439;

Best Local Similarity 100.0%; Pred. No. 2; 1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPGNGGLYSSG 15

|||||

Db 364 DRYPGNGGLYSSG 378

## RESULT 4

A59386

sanko - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-2001 #sequence\_revision 31-Dec-2001 #text\_change 31-Dec-2001

C:Accession: A59386

R:Sanko, S.

submitted to the Protein Sequence Database, March 2001

A:Reference number: A59386

A:Accession: A59386

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-2403 <SANK>

Query Match 52.3%; Score 46; DB 2; Length 2403;

Best Local Similarity 64.3%; Pred. No. 54;

Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG-LYSSG 15

|||||

Db 1752 PSSNGGFLFYASG 1765

## RESULT 5

S58870

reelin precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Dec-1996 #sequence\_revision 06-Dec-1996 #text\_change 11-Jan-2000

C:Accession: S58870; S71844; I49297

R:D'Arcangelo, G.; Milo, G.G.; Chen, S.C.; Soares, H.D.; Morgan, J.I.; Curran, T.

Nature 374, 719-723, 1995

A:Title: A protein related to extracellular matrix proteins deleted in the mouse mutant

A:Reference number: I49297; MUID:95231649; PMID:7715726

A:Accession: S58870

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3461 <DAR>

A:Cross-references: EMBL:U24703; NID:g902486; PID:g902487

R:D'Arcangelo, G.

submitted to the EMBL Data Library, April 1995

A:Reference number: S71844

A:Accession: S71844

A:Molecule type: mRNA

A:Residues: 1-215, 'T', 217-1905, 'S', 1907-3355, 'V', 3357-3391, 'N', 3393-3461 <DA2>

A:Cross-references: EMBL:U24703; NID:g902486; PID:g902487

C:Superfamily: unassigned EGF-related proteins; EGF homology

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-3461/Product: reelin #status predicted <MAT>

F:1769-1795/Domain: EGF homology <EGF>

Query Match 52.3%; Score 46; DB 2; Length 3461;

Best Local Similarity 71.4%; Pred. No. 76;

Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNGGLY--YSS 14

|||||

Db 1974 YPGNGIGLYCPYSS 1987

## RESULT 6

C82893

hypothetical protein U0425 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: C82893

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a

A:Reference number: A82870

A:Accession: C82893

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-225 <GLA>

A:Cross-references: GB:AE002139; GB:AF222894; NID:g6899405; PIDN:AAF30837.1; GSPDB:GN

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: U0425

A:Genetic code: SGC3

Query Match

48.9%; Score 43; DB 2; Length 225;

Best Local Similarity 70.0%; Pred. No. 18;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 YPSGNGGLY 12

|||||

Db 87 YPNGNAELYY 96

## RESULT 7

I49707

germ cell specific gene 3 protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 19-May-2000

C:Accession: I49707

R:Tanaka, H.; Yoshimura, Y.; Nishina, Y.; Nozaki, M.; Nojima, H.; Nishimura, Y.

FEBS Lett. 355, 4-10, 1994

A:Title: Isolation and characterization of cDNA clones specifically expressed in test

A:Reference number: I49707; MUID:95046372; PMID:7957958

A:Accession: I49707

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-284 <RES>

A:Cross-references: GB:P38545; NID:g603492; PID:g603493

C:Superfamily: actin-capping protein alpha chain

Query Match

48.9%; Score 43; DB 2; Length 284;

Best Local Similarity 60.0%; Pred. No. 22;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPGNGGL 10

|||||

Db 131 DHPNGNCNV 140

## RESULT 8

T24007  
 hypothetical protein R07B5.7 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: T24007  
 R:Kelly, P.  
 submitted to the EMBL Data Library, May 1996  
 A:Reference number: Z19829  
 A:Accession: T24007  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-339 <WIL>  
 A:Cross-references: EMBL:Z72512; PIDN:CAA96667.1; GSPDB:GN00023; CESP:R07B5.7  
 A:Experimental source: clone R07B5  
 C:Genetics:  
 A:Gene: CESP:R07B5.7  
 A:Map position: 5  
 A:Introns: 31/2; 93/1; 124/3; 194/3; 247/3; 267/3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C33A12.9b

Query Match 48.9%; Score 43; DB 2; Length 339;  
 Best Local Similarity 53.8%; Pred. No. 26;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYSSG 15  
 ||| :|||  
 Db 232 YPFGSVSVYSSG 244

RESULT 9  
 C86624  
 hypothetical protein CPj1070 [imported] - Chlamydomophila pneumoniae (strain J138)  
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
 C:Accession: C86624  
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, K.; Ouchi, F.; Shiba, T.; Is  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:20330349; PMID:10871362  
 A:Accession: C86624  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-456 <STO>  
 A:Cross-references: GB:BA000008; NID:g8979443; PIDN:BAA99277.1; GSPDB:GN00142  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: CPj1070

Query Match 48.9%; Score 43; DB 2; Length 456;  
 Best Local Similarity 60.0%; Pred. No. 34;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 YPSGNCGLY 12  
 :||| ||  
 Db 3 FPGCNCYY 12

RESULT 10  
 H72000  
 hypothetical protein CP0780 [imported] - Chlamydomophila pneumoniae (strains CWL029 and AR  
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
 C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
 C:Accession: H72000; A81538  
 R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
 Nature Genet. 21, 385-389, 1999  
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
 A:Reference number: A72000; MUID:99206606; PMID:10192388  
 A:Accession: H72000  
 A:Molecule type: DNA  
 A:Residues: 1-456 <ARN>  
 A:Cross-references: GB:AE001687; GB:AE001363; NID:g4377398; PIDN:AAI9207.1; PID:g437740  
 A:Experimental source: strain CWL029  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

, C.; Dodson, R.; Winn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jul-2000  
 C:Accession: A81538  
 R:Kelly, P.  
 submitted to the EMBL Data Library, May 1996  
 A:Reference number: Z19829  
 A:Accession: T24007  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-456 <REA>  
 A:Cross-references: GB:AE002238; GB:AE002161; NID:g7189693; PIDN:AAF38579.1; PID:g718  
 A:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 A:Gene: CPn1070; CP0780

Query Match 48.9%; Score 43; DB 2; Length 456;  
 Best Local Similarity 60.0%; Pred. No. 34;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 YPSGNCGLY 12  
 :||| ||  
 Db 3 FPGCNCYY 12

RESULT 11  
 G71829  
 probable outer membrane protein - Helicobacter pylori (strain J99)  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99  
 C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
 C:Accession: G71829; G71960  
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.  
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F  
 Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
 A:Reference number: A71800; MUID:99120557; PMID:9923682  
 A:Accession: G71829  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-696 <ARN>  
 A:Cross-references: GB:AE001549; GB:AE001439; NID:g4155858; PIDN:AAD06834.1; PID:g415  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: jhp0212

Query Match 48.3%; Score 42.5; DB 2; Length 696;  
 Best Local Similarity 69.2%; Pred. No. 61;  
 Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 3 YPSGNCGLYSSG 15  
 ||| ||| |||  
 Db 231 YPDGN-GNYSSG 242

RESULT 12  
 T16246  
 hypothetical protein F35A5.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000  
 C:Accession: T16246  
 R:Leimbach, D.  
 submitted to the EMBL Data Library, January 1996  
 A:Description: The sequence of C. elegans cosmid F35A5.  
 A:Reference number: Z18485  
 A:Accession: T16246  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-273 <LEI>  
 A:Cross-references: EMBL:U46675; NID:g1166613; PID:g1166614; PIDN:AA52645.1; GSPDB:G  
 A:Experimental source: strain Bristol N2; clone F35A5  
 C:Genetics:

A:Gene: CESP:F35A5.4  
A:Map position: X  
A:Introns: 18/3; 62/3; 81/3; 100/3; 140/3; 226/2

Query Match 46.6%; Score 41; DB 2; Length 273;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 PSNGCGLYSSG 15  
IIIIII II

DB 50 PSNGCGCGSG 61

## RESULT 13

A81289

hypothetical protein Cj1433c [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: A81289

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chillingworth, W.; Ouaifi, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell, A.; et al. (2001) The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: A81289

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <PAR>

A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73857.1; PID:g696886

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj1433c

C:Superfamily: Campylobacter jejuni hypothetical protein Cj1433c

Query Match 46.6%; Score 41; DB 2; Length 368;  
Best Local Similarity 70.0%; Pred. No. 58;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSNGCGLYVS 13  
IIIIII I.I

DB 265 PSNGCDNVHS 274

## RESULT 14

T26404

hypothetical protein Y105C5b.bb - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000

C:Accession: T26404

R:McMurray, A.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20208

A:Accession: T26404

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-385 <WIL>

A:Cross-references: EMBL:AL110479; PIDN:CAB54376.1; CESP:Y105C5b.bb

A:Experimental source: clone Y105C5b

C:Genetics:

A:Gene: CESP:Y105C5b.bb

A:Introns: 58/1; 112/1; 270/2

C:Superfamily: glutamate-ammonia ligase

Query Match 46.6%; Score 41; DB 2; Length 385;  
Best Local Similarity 53.3%; Pred. No. 61;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15  
IIIIII I.I

DB 344 DRRPSSGDPYTVTG 355

## RESULT 15

A72513

hypothetical protein APE2078 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: A72513

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.

DNA Res. 6; 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: A72513

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-437 <KAW>

A:Cross-references: DBJ:AF000063; NID:g5105654; PIDN:BAA81089.1; PID:g5105777

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2078

C:Superfamily: conserved hypothetical protein MTH1394

Query Match 46.6%; Score 41; DB 2; Length 437;

Best Local Similarity 57.1%; Pred. No. 69;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 RYPSGNCGLYSSG 15  
I.IIIIIII

DB 112 RYEGEAGLYLSSG 125

Search completed: December 2, 2002, 06:54:03

Job time: 3.05727 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:41:25 ; Search time 1.25551 Seconds  
(without alignments)  
495.533 Million cell updates/sec

Title: US-09-902-563-18  
Perfect score: 88  
Sequence: 1 DRYPSGNGLYSSG 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | %    | Description  |
|------------|-------|-------------|------|--------------|
| 1          | 88    | 100.0       | 432  | 1 FGL2_MOUSE |
| 2          | 88    | 100.0       | 439  | 1 FGL2_MOUSE |
| 3          | 46    | 52.3        | 3209 | 1 RELN_HUMAN |
| 4          | 46    | 52.3        | 3460 | 1 RELN_HUMAN |
| 5          | 46    | 52.3        | 3461 | 1 RELN_MOUSE |
| 6          | 46    | 52.3        | 3462 | 1 RELN_MOUSE |
| 7          | 45    | 51.1        | 210  | 1 CK17_HUMAN |
| 8          | 43    | 48.9        | 299  | 1 CAZ3_HUMAN |
| 9          | 43    | 48.9        | 299  | 1 CAZ3_MOUSE |
| 10         | 43    | 48.9        | 299  | 1 CAZ3_MOUSE |
| 11         | 41    | 46.6        | 121  | 1 CHA9_LYMDI |
| 12         | 41    | 46.6        | 437  | 1 YK78_AERPE |
| 13         | 41    | 46.6        | 517  | 1 DNB2_ADE07 |
| 14         | 41    | 46.6        | 590  | 1 VP02_BPRPD |
| 15         | 40    | 45.5        | 245  | 1 UNG_MYCGE  |
| 16         | 40    | 45.5        | 286  | 1 CAZ1_CHICK |
| 17         | 40    | 45.5        | 455  | 1 SYPC_METJA |
| 18         | 40    | 45.5        | 474  | 1 DNB2_ADE41 |
| 19         | 40    | 45.5        | 484  | 1 DNB2_ADE12 |
| 20         | 40    | 45.5        | 776  | 1 DNB7_MACFA |
| 21         | 40    | 45.5        | 1211 | 1 APTA_YEAST |
| 22         | 39.5  | 44.9        | 1391 | 1 LYS2_CANAL |
| 23         | 39    | 44.3        | 78   | 1 R27A_ASPOF |
| 24         | 39    | 44.3        | 79   | 1 R27A_HORVU |
| 25         | 39    | 44.3        | 79   | 1 R27A_MAIZE |
| 26         | 39    | 44.3        | 80   | 1 R27A_LYCES |
| 27         | 39    | 44.3        | 81   | 1 R27A_ARATH |
| 28         | 39    | 44.3        | 212  | 1 CK17_MOUSE |
| 29         | 39    | 44.3        | 420  | 1 Y461_SYNY3 |
| 30         | 39    | 44.3        | 673  | 1 SM4B_HUMAN |
| 31         | 39    | 44.3        | 754  | 1 AD07_HUMAN |
| 32         | 39    | 44.3        | 1077 | 1 XYNV_CLOTM |
| 33         | 39    | 44.3        | 1215 | 1 CLKA_BACTM |

|    |    |      |     |              |
|----|----|------|-----|--------------|
| 34 | 38 | 43.2 | 322 | 1 FSA_BRARE  |
| 35 | 38 | 43.2 | 329 | 1 HEM2_SCHFO |
| 36 | 38 | 43.2 | 365 | 1 GLM2_DROME |
| 37 | 38 | 43.2 | 367 | 1 GLNA_CAEEL |
| 38 | 38 | 43.2 | 373 | 1 GLNA_MOUSE |
| 39 | 38 | 43.2 | 373 | 1 GLNA_RAT   |
| 40 | 38 | 43.2 | 399 | 1 GLN1_DROME |
| 41 | 38 | 43.2 | 401 | 1 COAT_PAV   |
| 42 | 38 | 43.2 | 403 | 1 GLNA_SQUAC |
| 43 | 38 | 43.2 | 486 | 1 DNAB_RICPR |
| 44 | 38 | 43.2 | 803 | 1 Z226_HUMAN |
| 45 | 38 | 43.2 | 948 | 1 RPOP_PODAN |

## ALIGNMENTS

|   |  |
|---|--|
| RESULT 1  |  |
| FGL2_MOUSE  |  |
| ID FGL2_MOUSE STANDARD; PRT; 432 AA.  |  |
| AC P12804;  |  |
| DT 01-OCT-1989 (Rel. 12, Created)   |  |
| DT 01-OCT-1989 (Rel. 12, Last sequence update)                                |  |
| DT 16-OCT-2001 (Rel. 40, Last annotation update)                              |  |
| DE Fibrinogen precursor (Fibrinogen-like protein 2) (Prothrombinase)          |  |
| DE (Cytotoxic T-lymphocyte specific protein).                                 |  |
| GN FGL2 OR FIBLP.   |  |
| OS Mus musculus (Mouse).  |  |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;          |  |
| OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.         |  |
| OX NCBI_TaxID=10090;  |  |
| RN [1]  |  |
| RP SEQUENCE FROM N.A.   |  |
| RC TISSUE=Cytotoxic T-cell;   |  |
| RX MEDLINE=87175527; PubMed=3550794;  |  |
| RA Koyama T., Hall L.R., Hasegawa W.G., Toneygawa S., Saito H.;               |  |
| RT "Structure of a cytotoxic T-lymphocyte-specific gene shows a strong        |  |
| RT homology to fibrinogen beta and gamma chains.";                            |  |
| RL Proc. Natl. Acad. Sci. U.S.A. 84:1609-1613(1987).                          |  |
| RN [2]  |  |
| RP SEQUENCE FROM N.A. AND CHARACTERIZATION                                    |  |
| RC STRAIN-BALB/CJ; TISSUE=Peritoneal macrophage;                              |  |
| RX MEDLINE=95333285; PubMed=7609073;  |  |
| RA Parr R.L., Fung L., Reneker J., Myers-Mason N., Leibowitz J.L.,            |  |
| RA Levy G.;   |  |
| RT "Association of mouse fibrinogen-like protein with murine hepatitis        |  |
| RT virus-induced prothrombinase activity.";                                   |  |
| RL J. Virol. 69:5033-5038(1995).  |  |
| CC -!- FUNCTION: CONVERTS PROTHROMBIN TO THROMBIN.                            |  |
| CC -!- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED (BY SIMILARITY).               |  |
| CC -!- SUBCELLULAR LOCATION: Secreted   |  |
| CC -!- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN CYTOTOXIC              |  |
| CC T-CELLS.   |  |
| CC -!- INDUCTION: IN MACROPHAGES, DURING INFECTION BY MOUSE HEPATITIS         |  |
| CC VIRUS STRAIN 3 (MHV-3).  |  |
| CC -!- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.                   |  |
| CC This SWISS-PROT entry is copyright. It is produced through a collaboration |  |
| CC between the Swiss Institute of Bioinformatics and the EMBL outstation -    |  |
| CC the European Bioinformatics Institute. There are no restrictions on its    |  |
| CC use by non-profit institutions as long as its content is in no way         |  |
| CC modified and this statement is not removed. Usage by and for commercial    |  |
| CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ |  |
| CC or send an email to license@isb-sib.ch).                                   |  |
| CC -----  |  |
| CC EMBL; M16238; AAA37624.1; -  |  |
| CC EMBL; M15761; AAA37624.1; JOINED.  |  |
| CC EMBL; S78773; AAB34823.1; -  |  |
| CC PIR; A27447; A27447.   |  |
| CC HSP; P02671; IFZD.   |  |
| CC MGD; MGI:103266; Fgl2.   |  |
| CC InterPro: IPR002181; Fibrinogen_C.   |  |
| CC Pfam: PF00147; fibrinogen_C; 1.  |  |

DR SMART: SM00186; FBG; 1.  
 DR PROSITE: PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
 KW T-cell; Cytolysis; signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 432 FIBROLEUKIN.  
 FT DOMAIN 203 428 FIBRINOGEN C-TERMINAL.  
 FT DISULFID 206 235 BY SIMILARITY.  
 FT DISULFID 364 377 BY SIMILARITY.  
 FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 332 332 A -> G (IN REF. 2).  
 SQ SEQUENCE 432 AA; 48951 MW; 2B297F69CCBA4782 CRC64;

Query Match 100.0%; Score 88; DB 1; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 DRYPSGNGGLYSSG 15  
 |||||  
 357 DRYPSGNGGLYSSG 371

RESULT 2  
 FGL2\_HUMAN STANDARD; PRT; 439 AA.  
 AC Q14314;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Fibroleukin precursor (Fibrinogen-like protein 2) (p749).  
 GN FGL2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Small intestine;  
 RX MEDLINE=95369700; PubMed=7642106;  
 RA Ruegg C., Pytela R.;  
 RT "Sequence of a human transcript expressed in T-lymphocytes and  
 encoding a fibrinogen-like protein.";  
 RL Gene 160:257-262(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yuwaraj S., Liu M., Marsden P., Levy G.;  
 RT "Cloning and characterization of Hfg12: the human counterpart to the  
 mouse gene Fgl2.";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A. AND VARIANT GLU-53.  
 RA Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,  
 RA Poel C.L., Ozuna M., Yi Q., Nickerson D.A.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98309432; PubMed=9647217;  
 RA Marazzi S., Blum S., Hartmann R., Gundersen D., Schreyer M.,  
 RA Agraves S., von Fliedner V., Pytela R., Ruegg C.;  
 RT "Characterization of human fibroleukin, a fibrinogen-like protein  
 secreted by T lymphocytes.";  
 RL J. Immunol. 161:138-147(1998).  
 CC CC  
 CC -1- FUNCTION: MAY PLAY A ROLE IN PHYSIOLOGIC LYMPHOCYTE FUNCTIONS AT  
 CC MUCOSAL SITES.  
 CC -1- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN CYTOTOXIC  
 CC T-CELLS.  
 CC -1- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; Z36531; CAAS5298.1; .  
 DR EMBL; AF104015; AADI0825.1; .  
 DR EMBL; AF104014; AADI0825.1; JOINED.  
 DR EMBL; AF468959; AAL68855.1; .  
 DR HSP; P02671; 1PZD.  
 DR Genew; HGNC:3696; FGL2.  
 DR MIM; 605351; .  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR Pfam; PF00147; Fibrinogen\_C; 1.  
 DR SMART; SM00186; FBG; 1.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
 KW T-cell; Glycoprotein; Signal; Polymorphism.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 439 FIBROLEUKIN.  
 FT DOMAIN 210 435 FIBRINOGEN C-TERMINAL.  
 FT DISULFID 213 242 BY SIMILARITY.  
 FT DISULFID 371 384 BY SIMILARITY.  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 53 53 G -> E  
 FT SEQUENCE 439 AA; 50228 MW; DF34656288E49E68 CRC64;  
 SQ SEQUENCE 439 AA; 50228 MW; DF34656288E49E68 CRC64;  
 Query Match 100.0%; Score 88; DB 1; Length 439;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DRYPSGNGGLYSSG 15  
 |||||  
 364 DRYPSGNGGLYSSG 378

Db 364 DRYPSGNGGLYSSG 378

RESULT 3  
 RELN\_CHICK STANDARD; PRT; 3209 AA.  
 ID RELN\_CHICK  
 AC O93574;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Reelin (EC 3.4.21.-) (Fragment).  
 GN RELN.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bernier B., Goffinet A.M.;  
 RT "Comparative study of reelin in vertebrates.";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 CC CC  
 CC -1- FUNCTION: Extracellular matrix serine protease that plays a role  
 CC in layering of neurons in the cerebral cortex and cerebellum.  
 CC Regulates microtubule function in neurons and neuronal migration.  
 CC Affects migration of sympathetic preganglionic neurons in the  
 CC spinal cord, where it seems to act as a barrier to neuronal  
 CC migration. Enzymatic activity is important for the modulation of  
 CC cell adhesion. Binding to the extracellular domains of lipoprotein  
 CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of  
 CC Dab1 and modulation of Tau phosphorylation (By similarity).  
 CC -1- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By  
 CC similarity).





RA D'Arcangelo G., Miao G.G., Chen S.-C., Soares H.D., Morgan J.I.,  
RA Curran T.;  
RA "A protein related to extracellular matrix proteins deleted in the  
RA mouse mutant reeler";  
RA Nature 374:719-723(1995).  
RA [2]  
RA SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RA MEDLINE=98086481; PubMed=9417911;  
RA Royaux I., Lambert de Rouvroit C., D'Arcangelo G., Demirov D.,  
RA Goffinet A.M.;  
RA "Genomic organization of the mouse reeler gene";  
RA Genomics 46:240-250(1997).  
RA [3]  
RA SEQUENCE OF 2152-3461 FROM N.A. (ISOFORM 1).  
RA STRAIN=BAByC; TISSUE=Brain;  
RA MEDLINE=95375789; PubMed=7647795;  
RA Hirotsune S., Takahara T., Sasaki N., Hirose K., Yoshiki A.,  
RA Ohashi T., Kusakabe M., Murakami Y., Muramatsu M., Watanabe S.,  
RA Nakao K., Katsuki M., Hayashizaki Y.;  
RA "The reeler gene encodes a protein with an EGF-like motif expressed by  
RA pioneer neurons";  
RA Nat. Genet. 10:77-83(1995).  
RA [4]  
RA SEQUENCE OF 3044-3461 FROM N.A. (ISOFORM 2).  
RA STRAIN=C57Bl/6J; TISSUE=Testis;  
RA MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Schriml L.M., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Kuhl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Sakai K., Offord T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kameya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RA "Functional annotation of a full-length mouse cDNA collection";  
RA Nature 409:685-690(2001).  
RA [5]  
RA CHARACTERIZATION.  
RA MEDLINE=97141547; PubMed=8987733;  
RA D'Arcangelo G., Nakajima K., Miyata T., Ogawa M., Mikoshiba K.,  
RA Curran T.;  
RA "Reelin is a secreted glycoprotein recognized by the CR-50 monoclonal  
RA antibody";  
RA J. Neurosci. 17:23-31(1997).  
RA [6]  
RA CHARACTERIZATION.  
RA MEDLINE=21634904; PubMed=11689558;  
RA Quattrocchi C.C., Wannen F., Persico A.M., Ciafre S.A.,  
RA D'Arcangelo G., Parace M.G., Keller F.;  
RA "Reelin is a serine protease of the extracellular matrix";  
RA J. Biol. Chem. 277:303-309(2002).  
RA [7]  
RA TISSUE SPECIFICITY.  
RA MEDLINE=97325946; PubMed=9182958;  
RA Schiffrmann S.N., Bernier B., Goffinet A.M.;  
RA "Reelin mRNA expression during mouse brain development";  
RA Eur. J. Neurosci. 9:1055-1071(1997).  
RA [8]  
RA ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.  
RA MEDLINE=99263436; PubMed=10328932;  
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergoyck V.,  
RA Goffinet A.M.;  
RA "Evolutionarily conserved, alternative splicing of reelin during brain  
RA development";  
RA Exp. Neurol. 156:229-238(1999).  
RA [9]  
RA BINDING TO VLDL AND APOE2.  
RA MEDLINE=20036019; PubMed=10571241;  
RA Hiesberger T., Trzmsdorff M., Howell B.W., Goffinet A.M., Mumby M.C.,  
RA Cooper J.A., Herz J.;  
RA "Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces  
RA tyrosine phosphorylation of disabled-1 and modulates tau  
RA phosphorylation";  
RA Neuron 24:481-489(1999).  
RA [10]  
RA FUNCTION.  
RA MEDLINE=20359755; PubMed=10880573;  
RA Yip J.W., Yip Y.P.-L., Nakajima K., Capriotti C.;  
RA "Reelin controls position of autonomic neurons in the spinal cord";  
RA Proc. Natl. Acad. Sci. U.S.A. 97:8612-8616(2000).  
RA -1- FUNCTION: Extracellular matrix serine protease that plays a role  
RA in layering of neurons in the cerebral cortex and cerebellum.  
RA Regulates microtubule function in neurons and neuronal migration.  
RA Affects migration of sympathetic preganglionic neurons in the  
RA spinal cord, where it seems to act as a barrier to neuronal  
RA migration. Enzymatic activity is important for the modulation of  
RA cell adhesion. Binding to the extracellular domains of lipoprotein  
RA receptors VLDLR and ApoE2 induces tyrosine phosphorylation of  
RA Dab1 and modulation of tau phosphorylation.  
RA -1- SUBUNIT: Binds to the ectodomains of VLDLR and ApoE2.  
RA -1- SUBCELLULAR LOCATION: Secreted.  
RA -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are  
RA produced by alternative splicing.  
RA -1- TISSUE SPECIFICITY: The major isoform 1 is neuron-specific. It is  
RA abundantly produced during brain ontogenesis by the Cajal-Retzius  
RA cells and other pioneer neurons located in the telencephalic  
RA marginal zone and by granule cells of the external granular layer  
RA of the cerebellum. Expression is located in deeper layers in the  
RA developing hippocampus and olfactory bulb, low levels of  
RA expression are also detected in the immature striatum. At early  
RA developmental stages, expressed also in hypothalamic  
RA differentiation fields, tectum and spinal cord. A moderate to low  
RA level of expression occurs in the septal area, striatal fields,  
RA habenular nuclei, some thalamic nuclei, particularly the lateral  
RA geniculate, the retina and some nuclei of the reticular formation  
RA in the central field of the medulla. Very low levels found in  
RA liver and kidney. No expression in radial glial cells, cortical  
RA plate, Purkinje cells and inferior olivary neurons. The minor  
RA isoform 2 is only expressed in non neuronal cells. The minor  
RA isoform 3 is found in the same cells as isoform 1, but is almost  
RA undetectable in retina and brain stem.  
RA -1- DEVELOPMENTAL STAGE: First detected at embryonic day 11.5.  
RA Expression increases up to birth and remains high from post-natal  
RA day 2 to 11 in both cerebellum and forebrain. Expression  
RA declines thereafter and is largely brain specific in the adult.  
RA -1- DOMAIN: The basic C-terminal region is essential for secretion.  
RA -1- PTM: N-glycosylated and to a lesser extent also O-glycosylated.  
RA -1- DISEASE: Defects in rehl are the cause of the autosomal  
RA recessive reeler (rl) phenotype which is characterized by impaired  
RA motor coordination, tremors and ataxia. Neurons in affected mice  
RA fail to reach their correct locations in the developing brain,  
RA disrupting the organization of the cerebellar and cerebral  
RA cortices and other laminated regions.  
RA -1- SIMILARITY: BELONGS TO THE REELIN FAMILY.  
RA -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.  
RA -1- SIMILARITY: CONTAINS 15 BNR REPEATS.  
RA -----  
RA This SWISS-PROT entry is copyright. It is produced through a collaboration  
RA between the Swiss Institute of Bioinformatics and the EMBL Outstation  
RA the European Bioinformatics Institute. There are no restrictions on its  
RA use by non-profit institutions as long as its content is in no way  
RA modified and this statement is not removed. Usage by and for commercial  
RA entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
RA or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RA -----  
RA EMBL; U24703; AAB91599.1; "

EMBL; D63520; BAA09788.1; ALT\_INIT.  
 DR MGD; MGI:103022; Reel.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR002860; GH\_BNR.  
 DR InterPro; IPR002861; Reeler.  
 DR Pfam; PF00008; EGF; 5.  
 DR Pfam; PF02012; BNR; 15.  
 DR Pfam; PF02014; Reeler; 1.  
 DR SMART; SM00181; EGF; 5.  
 DR SMART; SM00001; EGF-like; 2.  
 DR PROSITE; PS00022; EGF-1; 7.  
 DR PROSITE; PS01186; EGF-2; 6.  
 DR Hydrolase; Serine protease; Developmental protein; Matrix protein;  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;  
 KW Alternative splicing.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 3461 REELIN.  
 FT DOMAIN 40 172 REELER.  
 FT DOMAIN 671 702 EGF-LIKE 1.  
 FT DOMAIN 1030 1061 EGF-LIKE 2.  
 FT DOMAIN 1409 1442 EGF-LIKE 3.  
 FT DOMAIN 1765 1796 EGF-LIKE 4.  
 FT DOMAIN 2129 2161 EGF-LIKE 5.  
 FT DOMAIN 2478 2509 EGF-LIKE 6.  
 FT DOMAIN 2853 2884 EGF-LIKE 7.  
 FT DOMAIN 3228 3260 EGF-LIKE 8.  
 FT REPEAT 593 604 BNR 1.  
 FT REPEAT 799 810 BNR 2.  
 FT REPEAT 952 963 BNR 3.  
 FT REPEAT 1157 1168 BNR 4.  
 FT REPEAT 1323 1334 BNR 5.  
 FT REPEAT 1535 1546 BNR 6.  
 FT REPEAT 1686 1697 BNR 7.  
 FT REPEAT 1884 1895 BNR 8.  
 FT REPEAT 2043 2054 BNR 9.  
 FT REPEAT 2250 2261 BNR 10.  
 FT REPEAT 2399 2410 BNR 11.  
 FT REPEAT 2598 2609 BNR 12.  
 Query Match 52.3%; Score 46; DB 1; Length 3461;  
 Best Local Similarity 71.4%; Pred. No. 40;  
 Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;  
 Qy 3 YFSGNGLY--YSS 14  
 || || || || ||  
 Db 1974 YFGNIGLYCPYSS 1987  
 RESULT 6  
 RELN\_RAT  
 ID RELN\_RAT STANDARD; PRT; 3462 AA.  
 AC P58751;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Reelin precursor (EC 3.4.21.-).  
 GN RELN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 [1]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Cerebellum;  
 RA Kikkawa S., Terashima T.;  
 RL Submitted (Sep-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN ALTERNATIVE SPLICING.  
 RP MBLIN-99263436; PubMed=10328932;  
 RX Lambert de Rouvroit C., Bernier B., Royaux I., de Bergueyck V.,  
 RA Coffinet A.M.;  
 FT "Evolutionarily conserved, alternative splicing of reelin during brain

development".  
 RL Exp. Neurol. 156:229-238(1999).  
 CC -!- FUNCTION: Extracellular matrix serine protease that plays a role  
 CC in layering of neurons in the cerebral cortex and cerebellum.  
 CC Regulates microtubule function in neurons and neuronal migration.  
 CC Affects migration of sympathetic preganglionic neurons in the  
 CC spinal cord, where it seems to act as a barrier to neuronal  
 CC migration. Enzymatic activity is important for the modulation of  
 CC cell adhesion. Binding to the extracellular domains of lipoprotein  
 CC receptors VLDL and ApoE2 induces tyrosine phosphorylation of  
 CC Dab1 and modulation of Tau phosphorylation (By similarity).  
 CC -!- SUBUNIT: Binds to the ectodomains of VLDL and ApoE2 (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are  
 CC produced by alternative splicing.  
 CC -!- TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis  
 CC by the Cajal-Retzius cells and other pioneer neurons located in  
 CC the telencephalic marginal zone and by granule cells of the  
 CC external granular layer of the cerebellum.  
 CC -!- DOMAIN: The basic C-terminal region is essential for secretion (By  
 CC similarity).  
 CC -!- SIMILARITY: BELONGS TO THE REELIN FAMILY.  
 CC -!- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 15 BNR REPEATS.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 -----  
 DR EMBL; AB049473; BAB78470.1; -.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR002860; GH\_BNR.  
 DR InterPro; IPR002861; Reeler.  
 DR Pfam; PF00008; EGF; 5.  
 DR Pfam; PF02012; BNR; 15.  
 DR Pfam; PF02014; Reeler; 1.  
 DR PROSITE; PS00022; EGF-1; 7.  
 DR PROSITE; PS01186; EGF-2; 6.  
 KW Hydrolase; Serine protease; Developmental protein; Matrix protein;  
 KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;  
 KW Alternative splicing.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 3462 REELIN.  
 FT DOMAIN 41 173 REELER.  
 FT DOMAIN 672 703 EGF-LIKE 1.  
 FT DOMAIN 1031 1062 EGF-LIKE 2.  
 FT DOMAIN 1410 1443 EGF-LIKE 3.  
 FT DOMAIN 1766 1797 EGF-LIKE 4.  
 FT DOMAIN 2130 2162 EGF-LIKE 5.  
 FT DOMAIN 2479 2510 EGF-LIKE 6.  
 FT DOMAIN 2854 2885 EGF-LIKE 7.  
 FT DOMAIN 3229 3261 EGF-LIKE 8.  
 FT REPEAT 594 605 BNR 1.  
 FT REPEAT 800 811 BNR 2.  
 FT REPEAT 953 964 BNR 3.  
 FT REPEAT 1158 1169 BNR 4.  
 FT REPEAT 1324 1335 BNR 5.  
 FT REPEAT 1536 1547 BNR 6.  
 FT REPEAT 1687 1698 BNR 7.  
 FT REPEAT 1885 1896 BNR 8.  
 FT REPEAT 2044 2055 BNR 9.  
 FT REPEAT 2251 2262 BNR 10.  
 FT REPEAT 2400 2411 BNR 11.  
 FT REPEAT 2599 2610 BNR 12.  
 FT REPEAT 2779 2790 BNR 13.  
 FT REPEAT 2980 2991 BNR 14.  
 FT REPEAT 3364 3375 BNR 15.  
 FT DOMAIN 3433 3462 ARG-RICH (BASIC).

```

FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1448 1448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1601 1601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1751 1751 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1922 1922 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2146 2146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2270 2270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2318 2318 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2570 2570 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2963 2963 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3017 3017 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3074 3074 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3186 3186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3413 3413 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3440 3440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 3430 3431 MISSING (IN ISOFORM 2).
FT VARSPLIC 3430 3462 MISSING (IN ISOFORM 3).
SEQUENCE 3462 AA; 387525 MW; PCCF89B090E035F6 CRC64;

Query Match 52.38; Score 46; DB 1; Length 3462;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNGCLY--YSS 14
||| ||| |||
Db 1975 YPGNGICLYCPYSS 1988

RESULT 7
CK17_HUMAN STANDARD; PRT; 210 AA.
AC Q9N031;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein C11orf17.
GN C11orf17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=21418998; PubMed=11528127;
Anid C., Bahr A., Mujica A., Sampson N., Bikar S.E., Winterpacht A.,
Zabel B., Hankeln T., Schmidt E.K.;
RT "Comparative genomic sequencing reveals a strikingly similar
architecture of a conserved syntenic region on human chromosome
11p15.3 (including gene SRS) and mouse chromosome 7.";
RL Cytogenet. Cell Genet. 93:284-290(2001).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DB EMBL; AJ400877; C11orf17.
Genew; HGNC:1170; C11orf17.
SEQUENCE 210 AA; 23086 MW; CA3AB6CF4ECD5E59 CRC64;

Query Match 51.1%; Score 45; DB 1; Length 210;
Best Local Similarity 66.7%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNGCLYSS 14

```

```

Db 95 YTSSQCKYISS 106
||| ||| ||| |||
RESULT 8
CAZ3_HUMAN STANDARD; PRT; 299 AA.
ID CAZ3_HUMAN STANDARD; PRT; 299 AA.
AC Q96KX2; O969J0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE F-actin capping protein alpha-3 subunit (Cap2 alpha-3) (GSG3).
GN CAP2A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
TISSUE=Testis;
RA Miyagawa Y., Tanaka H.;
RT "Cloning and characterization of human actin capping protein, Gsg3.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
TISSUE=Testis;
RA Kawakami B., Sugiyama A., Takemoto M., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Nishikawa T., Nagai K., Isogai T.,
RA Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: F-ACTIN CAPPING PROTEINS BIND IN A CA(2+)-INDEPENDENT
MANNER TO THE FAST GROWING ENDS OF ACTIN FILAMENTS (BARBED END)
THEREBY BLOCKING THE EXCHANGE OF SUBUNITS AT THESE ENDS. UNLIKE
OTHER CAPPING PROTEINS (SUCH AS GELSOLIN AND SEVERIN), THESE
PROTEINS DO NOT SEVER ACTIN FILAMENTS. MAY PLAY A ROLE IN THE
MORPHOGENESIS OF SPERMATID (BY SIMILARITY).
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE F-ACTIN CAPPING PROTEIN ALPHA SUBUNIT
FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DB EMBL; AB053259; BAB61901.1;
EMBL; BC016745; AAH16745.1;
EMBL; AK058174; BAB71703.1;
InterPro: IPR002189; F-actin_cap_A.
Pfam: PF01267; F-actin_cap_A; 1.
DR PRODOM: PD006960; F-actin_cap_A; 1.
DR PROSITE; PS00748; F-ACTIN_CAPPING_A.1; FALSE_NEG.
DR PROSITE; PS00749; F-ACTIN_CAPPING_A.2; 1.
KW Actin-binding; Capping protein; Multigene family.
FT CONFLICT 95 95 D -> Y (IN REF. 2).
SQ SEQUENCE 299 AA; 35024 MW; CEE130C8F7397F67 CRC64;

Query Match 48.9%; Score 43; DB 1; Length 299;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DRYPSGNGCL 10.

```

Db 131 DHPGNCNM 140

RESULT 9  
CAZ3\_MOUSE  
ID CAZ3\_MOUSE STANDARD; PRT; 299 AA.  
AC P70190;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE F-actin capping protein alpha-3 subunit (Cap2 alpha-3) (Germ cell-specific protein 3).  
GN CAP2A3 OR CAPPA3 OR GSG3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6; TISSUE=Testis;  
RA Tanaka H., Yoshimura Y., Nishina Y., Nozaki M., Nojima H., Nishimune Y.;  
RT "Isolation and characterization of cDNA clones specifically expressed in testicular germ cells.";  
RL FEBS Lett. 355:4-10(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Yoshimura Y., Tanaka H., Nozaki M., Yomogida K., Shimamura K., Yasunaga T., Nishimune Y.;  
RT "Genomic analysis of male germ cell-specific actin capping protein alpha.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: F-ACTIN CAPPING PROTEINS BIND IN A CA(2+)-INDEPENDENT MANNER TO THE FAST GROWING ENDS OF ACTIN FILAMENTS (BARBED END) THEREBY BLOCKING THE EXCHANGE OF SUBUNITS AT THESE ENDS. UNLIKE OTHER CAPPING PROTEINS (SUCH AS GELSOLIN AND SEVERIN), THESE PROTEINS DO NOT SEVER ACTIN FILAMENTS. MAY PLAY A ROLE IN THE MORPHOGENESIS OF SPERMATID.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: EXCLUSIVELY EXPRESSED IN THE TESTIS.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN 24-DAY-OLD AND ADULT TESTIS, BUT NOT IN 4-, 10- AND 16-DAY-OLD TESTIS.  
CC -1- SIMILARITY: BELONGS TO THE F-ACTIN CAPPING PROTEIN ALPHA SUBUNIT FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: D87471; BAA13409.1; -  
EMBL: AB026984; BAA81887.1; -  
MGD: MGI:106221; Cappa3.  
InterPro: IPR002189; F-actin\_cap\_A.  
Pfam: PF01267; F-actin\_cap\_A; 1.  
ProDom: PD006960; F-actin\_cap\_A; 1.  
PROSITE: PS00748; F-ACTIN\_CAPPING\_A.1; FALSE\_NEG.  
DR PROSITE: PS00749; F-ACTIN\_CAPPING\_A.2; 1.  
KW Actin-binding; Capping protein; Multigene family.  
SQ SEQUENCE 299 AA; 34952 MW; 8C875793F3233C6 CRC64;

Query Match 48.9%; Score 43; DB 1; Length 299;  
Best Local Similarity 60.0%; Pred. No. 11;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPSGNCGL 10  
| | | | |

Db 131 DHPGNCNV 140

RESULT 10  
CAZ3\_RAT  
ID CAZ3\_RAT STANDARD; PRT; 299 AA.  
AC Q9WUV6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE F-actin capping protein alpha-3 subunit (Cap2 alpha-3).  
GN CAP2A3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=98069272; PubMed=9406198;  
RA Hurst S., Howes E.A., Coadwell J., Jones R.;  
RT "Expression of a testis-specific putative actin-capping protein associated with the developing acrosome during rat spermiogenesis.";  
RL Mol. Reprod. Dev. 49:81-91(1998).  
CC -1- FUNCTION: F-ACTIN CAPPING PROTEINS BIND IN A CA(2+)-INDEPENDENT MANNER TO THE FAST GROWING ENDS OF ACTIN FILAMENTS (BARBED END) THEREBY BLOCKING THE EXCHANGE OF SUBUNITS AT THESE ENDS. UNLIKE OTHER CAPPING PROTEINS (SUCH AS GELSOLIN AND SEVERIN), THESE PROTEINS DO NOT SEVER ACTIN FILAMENTS. MAY PLAY A ROLE IN THE MORPHOGENESIS OF SPERMATID.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: EXCLUSIVELY EXPRESSED IN THE TESTIS.  
CC -1- SIMILARITY: BELONGS TO THE F-ACTIN CAPPING PROTEIN ALPHA SUBUNIT FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: Y12538; CAA73137.1; -  
InterPro: IPR002189; F-actin\_cap\_A.  
Pfam: PF01267; F-actin\_cap\_A; 1.  
PRINTS: PR00191; FACTINCAPA.  
ProDom: PD006960; F-actin\_cap\_A; 1.  
PROSITE: PS00748; F-ACTIN\_CAPPING\_A.1; FALSE\_NEG.  
DR PROSITE: PS00749; F-ACTIN\_CAPPING\_A.2; 1.  
KW Actin-binding; Capping protein; Multigene family.  
SQ SEQUENCE 299 AA; 35007 MW; 3D753088BCF79BE5 CRC64;

Query Match 48.9%; Score 43; DB 1; Length 299;  
Best Local Similarity 60.0%; Pred. No. 11;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPSGNCGL 10  
| | | | |

Db 131 DHPGNCNV 140

RESULT 11  
CHA9\_LYMDI  
ID CHA9\_LYMDI STANDARD; PRT; 121 AA.  
AC P50603;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Chorion class A proteins LD9 (Fragment).  
OS Lymantia dispar (Gypsy moth).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;  
 OC Ditrysia; Noctuoidea; Lymantriidae; Lymantria.  
 OX NCBI\_TaxID=13123;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Chorion follicles;  
 RX MEDLINE=95018300; PubMed=7932786;  
 RA Leclerc R.F., Regier J.C.;  
 RT "Evolution of chorion gene families in lepidoptera: characterization  
 of 15 cDNAs from the gypsy moth.";  
 RL J. Mol. Evol. 39:244-254(1994).  
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE  
 CC GYPSY MOTH.  
 CC -!- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH  
 CC BELONG CLASSES A, CA AND HCA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U04661; AAA67861.1; -  
 KW Eggshell; Chorion; Repeat; Multigene family.  
 FT NON-TER 1  
 SQ SEQUENCE 121 AA; 11617 MW; 528E28501ED06617 CRC64;  
 Query Match 46.6%; Score 41; DB 1; Length 121;  
 Best Local Similarity 66.7%; Pred. No. 9.5;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 RYPSGNCGL 10  
 DB 41 RYPAGACGI 49  
 RESULT 12  
 YK78\_AERPE  
 ID YK78\_AERPE STANDARD; PRT; 437 AA.  
 AC Q9YA60;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein APE2078.  
 GN APE2078.  
 OS Aeropyrum pernix.  
 RA Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
 RA Desulfurococaceae; Aeropyrum.  
 NCBI\_TaxID=56636;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,  
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 RT crenarchaeon, Aeropyrum pernix K1.";  
 RL DNA Res. 6:83-101(1999).  
 CC -!- SIMILARITY: BELONGS TO THE URID FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: AP000063; BA881089.1; -  
 DR InterPro: IPR002830; carboxylase.  
 DR Pfam: PF01977; UPF0096; 1.  
 DR TIGRfams: TIGR00148; UPF0096; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 437 AA; 47635 MW; 5055140A0A2D602 CRC64;  
 Query Match 46.6%; Score 41; DB 1; Length 437;  
 Best Local Similarity 57.1%; Pred. No. 33;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 RYPSGNCGLYSSG 15  
 DB 112 RYEGEAGLYLSSG 125  
 RESULT 13  
 DNE2\_ADE07  
 ID DNE2\_ADE07 STANDARD; PRT; 517 AA.  
 AC P04497;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DE 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Early E2A DNA-binding protein.  
 GN DBP.  
 OS Human adenovirus type 7.  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=10519;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84185604; PubMed=6325415;  
 RA Quinn C.O., Kitchingman G.R.;  
 RT "Sequence of the DNA-binding protein gene of a human subgroup B  
 RT adenovirus (type 7). Comparisons with subgroup C (type 5) and  
 RT subgroup A (type 12).";  
 RL J. Biol. Chem. 259:5003-5009(1984).  
 CC -!- FUNCTION: BINDS COOPERATIVELY SINGLE-STRANDED DNA IN A SEQUENCE-  
 CC INDEPENDENT MANNER. INVOLVED IN DNA-REPLICATION, REGULATION OF  
 CC MRNA FORMATION, AND HOST-RANGE SPECIFICITY. ZINC IS REQUIRED  
 CC FOR DNA BINDING.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR. ACCUMULATES IN INFECTED CELLS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: K02530; AAA42508.1; -  
 DR PIR: A03834; ERADA7.  
 DR HSP; P03265; IADU.  
 DR InterPro: IPR005376; Vir\_DNA\_Zn\_bind.  
 DR InterPro: IPR003176; Vir\_DNA\_binding.  
 DR Pfam: PF02236; Vir\_DNA\_binding; 1.  
 DR Pfam: PF03728; Vir\_DNA\_Zn\_bind; 2.  
 KW Early protein; DNA-binding; zinc-finger; Phosphorylation;  
 KW Nuclear protein.  
 FT MOD\_RES 180 180 PHOSPHORYLATION (PROBABLE).  
 FT ZN\_FING 258 271 POTENTIAL.  
 SQ SEQUENCE 517 AA; 58306 MW; A2CC8B1CA1F191F CRC64;  
 Query Match 46.6%; Score 41; DB 1; Length 517;  
 Best Local Similarity 33.3%; Pred. No. 38;  
 Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 DRYPGNCGLYSSG 15  
 DB 333 NOFSSKSGMEYTEG 347

```

RESULT 14
VP02_BPPRD STANDARD; PRT; 590 AA.
AC P27378;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE 01-DEC-1992 (Rel. 24, Last annotation update)
DE Adsorption protein P2.
GN II.
OS Bacteriophage PRD1.
OC Viruses; dsDNA viruses, no RNA stage; Tectiviridae; Tectivirus.
OX NCBI_TaxID=10658;
RN [1].
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-9.
RX MEDLINE=91306449; PubMed=1853567;
RA Bamford J.K.H., Haeninen A.-L., Pakula T.M., Ojala P.M.,
RA Kaikkinen N., Frilander M., Bamford D.H.;
RT "Genome organization of membrane-containing bacteriophage PRD1.";
RL Virology 183:658-676(1991).
CC -!- FUNCTION: THE LIFE CYCLE OF THE PHAGE BEGINS WHEN THE PHAGE
CC ADSORBS TO ITS RECEPTOR ON THE SURFACE OF THE HOST VIA THE
CC ADSORPTION PROTEIN P2 AND INJECTS ITS DNA INTO THE HOST CYTOPLASM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M69077; AAA32458.1; -
DR PIR; D40477; WMBPQ2.
FT INIT_MET 0
SQ SEQUENCE 590 AA; 63690 MW; 5ACA024C92B8F6BC CRC64;

Query Match 46.6%; Score 41; DB 1; Length 590;
Best Local Similarity 46.7%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 DRYPNGNGLYSSG 15
   I : : : : : I I
Db 407 DEWYANNGLFPMSC 421

RESULT 15
UNG_MYCSE
ID UNG_MYCSE STANDARD; PRT; 245 AA.
RX P47343; Q49318;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-FEB-1996 (Rel. 33, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Uracil-DNA glycosylase (EC 3.2.2.-) (UDG).
GN UNG OR MG097.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1].
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uitterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2].
RP SEQUENCE OF 74-231 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;

```

```

RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -!- FUNCTION: EXCISES URACIL RESIDUES FROM THE DNA WHICH CAN ARISE
CC AS A RESULT OF MISINCORPORATION OF DUMP RESIDUES BY DNA
CC POLYMERASE OR DUE TO DEAMINATION OF CYTOSINE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE URACIL-DNA GLYCOSYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39690; AAC71315.1; -
DR EMBL; U02201; AAD12490.1; -
DR HSP; P12295; LEUG.
DR TIGR; MG097; -
DR InterPro; IPR005122; UDNA_glycos.
DR InterPro; IPR002043; U_DNA_glycosyl.
DR InterPro; IPR003249; U_glycosyl.
DR Pfam; PF03167; UDG; 1.
DR ProDom; PD001589; U_glycosyl; 1.
DR TIGRFAMs; TIGR00628; ung; 1.
DR PROSITE; PS00130; U_DNA_GLYCOSYLASE; 1.
KW DNA repair; Hydrolase; Glycosidase; Complete proteome.
FT ACT_SITE 72 72 GENERAL BASE (BY SIMILARITY).
FT CONFLICT 193 205 EHLSPHPSPLSG -> GAFIISPLTKC (IN
FT REF. 2).
SQ SEQUENCE 245 AA; 28188 MW; AZEDA605D83D7CCC CRC64;

Query Match 45.5%; Score 40; DB 1; Length 245;
Best Local Similarity 56.2%; Pred. No. 27;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 1 DRYPNGN--CGLYSS 14
   I : : : : I I I : : I
Db 72 DPYSPNDACGLAFAS 87

```

Search completed: December 2, 2002, 06:51:13  
Job time : 3.25551 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:48:20 ; Search time 4.09692 Seconds  
(without alignments)  
754.398 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_prodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID        | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 88    | 100.0       | 357    | 11 Q9EPT7 | Q9ept7 rattus norv |
| 2          | 56    | 63.6        | 220    | 5 Q8T8A2  | Q8t8a2 ciona savig |
| 3          | 56    | 63.6        | 652    | 5 Q9NDQ1  | Q9ndq1 ciona intes |
| 4          | 46    | 52.3        | 1785   | 4 Q9Y211  | Q9y211 homo sapien |
| 5          | 46    | 52.3        | 1785   | 4 Q9Y4V9  | Q9y4v9 homo sapien |
| 6          | 46    | 52.3        | 2403   | 4 Q9UGM2  | Q9ugm2 homo sapien |
| 7          | 46    | 52.3        | 2412   | 4 Q9UJ57  | Q9uj57 homo sapien |
| 8          | 46    | 52.3        | 2413   | 4 Q9UKJ4  | Q9ukj4 homo sapien |
| 9          | 46    | 52.3        | 2413   | 4 Q96DU4  | Q96du4 homo sapien |
| 10         | 46    | 52.3        | 2426   | 4 Q9UGM3  | Q9ugm3 homo sapien |
| 11         | 45    | 51.1        | 183    | 4 Q8TAE0  | Q8tae0 homo sapien |
| 12         | 45    | 51.1        | 210    | 4 Q8TAC6  | Q8tac6 homo sapien |
| 13         | 44.5  | 50.6        | 279    | 11 Q9CU50 | Q9cu50 mus musculu |
| 14         | 44.5  | 50.6        | 457    | 11 Q8R0Z6 | Q8r0z6 mus musculu |
| 15         | 44    | 50.0        | 217    | 7 Q9CJH7  | Q9cjh7 salmo trutt |
| 16         | 44    | 50.0        | 259    | 5 Q764E8  | Q764e8 cryptospori |

|    |      |      |      |           |                     |
|----|------|------|------|-----------|---------------------|
| 17 | 44   | 50.0 | 292  | 5 Q9VDM9  | Q9vdm9 drosophila   |
| 18 | 44   | 50.0 | 472  | 17 Q8TZD1 | Q8tzd1 pyrococcus   |
| 19 | 44   | 50.0 | 544  | 5 Q8SW93  | Q8sw93 encephalito  |
| 20 | 43   | 48.9 | 225  | 16 Q9PG63 | Q9pg63 ureaplasma   |
| 21 | 43   | 48.9 | 299  | 11 Q9D4N3 | Q9d4n3 mus musculu  |
| 22 | 43   | 48.9 | 339  | 5 Q21788  | Q21788 caenorhabdi  |
| 23 | 43   | 48.9 | 456  | 16 Q9Z619 | Q9z619 chlamydia p  |
| 24 | 43   | 48.9 | 1007 | 10 Q8VYR8 | Q8vyr8 arabidopsis  |
| 25 | 43   | 48.9 | 1082 | 10 Q8W0U0 | Q8w0u0 sorghum bic  |
| 26 | 43   | 48.9 | 2862 | 16 Q8R874 | Q8r874 thermopaneer |
| 27 | 42.5 | 48.3 | 178  | 2 Q87568  | Q87568 helicobacte  |
| 28 | 42.5 | 48.3 | 696  | 16 Q9Z390 | Q9z390 helicobacte  |
| 29 | 42   | 47.7 | 664  | 5 Q16977  | Q16977 caenorhabdi  |
| 30 | 42   | 47.7 | 771  | 5 Q9NGD0  | Q9ngd0 leishmania   |
| 31 | 42   | 47.7 | 771  | 5 Q8T7E0  | Q8t7e0 leishmania   |
| 32 | 42   | 47.7 | 873  | 5 Q8WTL6  | Q8wtl6 caenorhabdi  |
| 33 | 41.5 | 47.2 | 470  | 4 Q9BZ20  | Q9bzz0 homo sapien  |
| 34 | 41   | 46.6 | 154  | 10 Q9FF72 | Q9ff72 arabidopsis  |
| 35 | 41   | 46.6 | 273  | 5 Q20000  | Q20000 caenorhabdi  |
| 36 | 41   | 46.6 | 351  | 10 Q8W006 | Q8w006 gelidium cr  |
| 37 | 41   | 46.6 | 368  | 16 Q9PMU4 | Q9pmu4 campylobact  |
| 38 | 41   | 46.6 | 388  | 5 Q9U307  | Q9u307 caenorhabdi  |
| 39 | 41   | 46.6 | 392  | 2 Q9KX40  | Q9kx40 pseudomonas  |
| 40 | 41   | 46.6 | 457  | 12 Q9W972 | Q9w972 porcine ade  |
| 41 | 41   | 46.6 | 464  | 4 Q961C9  | Q961c9 homo sapien  |
| 42 | 41   | 46.6 | 480  | 12 Q8QUK8 | Q8quk8 infectious   |
| 43 | 41   | 46.6 | 488  | 4 Q96NM4  | Q96nm4 homo sapien  |
| 44 | 41   | 46.6 | 634  | 5 Q96LJ9  | Q96lj9 drosophila   |
| 45 | 41   | 46.6 | 637  | 5 Q9VDJ6  | Q9vdj6 drosophila   |

## ALIGNMENTS

### RESULT 1

Q9EPT7 PRELIMINARY; PRT; 357 AA.  
ID Q9EPT7  
AC Q9EPT7;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Prothrombinase FGL2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RA Rychlik D.F., Chien E., Philippe M.;  
RT FGL2 expression in the Sprague-Dawley Rat.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DDJB databases.  
DR EMBL; AF323608; AAC42269.1; -  
DR HSSP; P02671; IFZD.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR Pfam; PF00147; fibrinogen\_C; 1.  
DR SMART; SM00186; FSG; 1.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
SQ SEQUENCE 357 AA; 40966 MW; 31183DD9A02EBBA9 CRC64;

Query Match 100.0%; Score 88; DB 11; Length 357;  
Best Local Similarity 100.0%; Pred. No. 8e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRYPSGNCGLYSSG 15  
Db 282 DRYPSGNCGLYSSG 296

### RESULT 2

Q8T8A2 PRELIMINARY; PRT; 220 AA.  
ID Q8T8A2  
AC Q8T8A2;

```

DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fibrinogen-like protein (Fragment).
GN CS-FIBRINOGEN-LIKE.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920613; PubMed=11933208;
RA Imai K.S., Satoh N., Satou Y.;
RT "Early embryonic expression of FGF4/6/9 gene and its role in the
RT induction of mesenchyme and notochord in Ciona savignyi embryos.";
RL Development 129:1729-1738(2002).
DR EMBL; AB073373; BAB86674.1; -.
FT NON_TER 1
SQ SEQUENCE 220 AA; 25758 MW; 4160FAE727F3ED06 CRC64;

Query Match 63.6%; Score 56; DB 5; Length 220;
Best Local Similarity 66.7%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
DB 143 DGYALGNGCRYRSG 157

RESULT 3
Q9NDQ1 PRELIMINARY; PRT; 652 AA.
AC Q9NDQ1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Fibrinogen-like protein.
GN CI-FIBRN.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RX Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,
RA Satoh N.;
RT "Characterization of Brachyury downstream notochord genes in the Ciona
RT intestinalis embryo.";
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036849; BAB00626.1; -.
DR HSSP; P02671; 1FZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRN_AG_C_DOMAIN; 1.
SQ SEQUENCE 652 AA; 73252 MW; A452BA325162F0E0 CRC64;

Query Match 63.6%; Score 56; DB 5; Length 652;
Best Local Similarity 66.7%; Pred. No. 0.39;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
DB 574 DGYALGNGCRYRSG 588

RESULT 4
Q9Y211 PRELIMINARY; PRT; 1785 AA.
AC Q9Y211;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

```

```

DE DMBT1.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20017478; PubMed=10551316;
RA Takeshita H., Sato M., Shiwaku H.O., Semba S., Sakurada A., Hoshi M.,
RA Hayashi Y., Tagawa Y., Ayabe H., Horii A.;
RT "Expression of the DMBT1 gene is frequently suppressed in human lung
RT cancer.";
RL Jpn. J. Cancer Res. 90:903-908(1999).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AB020851; BAA78577.1; JOINED.
DR EMBL; AB020812; BAA78577.1; JOINED.
DR EMBL; AB020813; BAA78577.1; JOINED.
DR EMBL; AB020814; BAA78577.1; JOINED.
DR EMBL; AB020815; BAA78577.1; JOINED.
DR EMBL; AB020816; BAA78577.1; JOINED.
DR EMBL; AB020817; BAA78577.1; JOINED.
DR EMBL; AB020818; BAA78577.1; JOINED.
DR EMBL; AB020819; BAA78577.1; JOINED.
DR EMBL; AB020820; BAA78577.1; JOINED.
DR EMBL; AB020821; BAA78577.1; JOINED.
DR EMBL; AB020822; BAA78577.1; JOINED.
DR EMBL; AB020823; BAA78577.1; JOINED.
DR EMBL; AB020824; BAA78577.1; JOINED.
DR EMBL; AB020825; BAA78577.1; JOINED.
DR EMBL; AB020826; BAA78577.1; JOINED.
DR EMBL; AB020827; BAA78577.1; JOINED.
DR EMBL; AB020828; BAA78577.1; JOINED.
DR EMBL; AB020829; BAA78577.1; JOINED.
DR EMBL; AB020830; BAA78577.1; JOINED.
DR EMBL; AB020831; BAA78577.1; JOINED.
DR EMBL; AB020832; BAA78577.1; JOINED.
DR EMBL; AB020833; BAA78577.1; JOINED.
DR EMBL; AB020834; BAA78577.1; JOINED.
DR EMBL; AB020835; BAA78577.1; JOINED.
DR EMBL; AB020836; BAA78577.1; JOINED.
DR EMBL; AB020837; BAA78577.1; JOINED.
DR EMBL; AB020838; BAA78577.1; JOINED.
DR EMBL; AB020839; BAA78577.1; JOINED.
DR EMBL; AB020840; BAA78577.1; JOINED.
DR EMBL; AB020841; BAA78577.1; JOINED.
DR EMBL; AB020842; BAA78577.1; JOINED.
DR EMBL; AB020843; BAA78577.1; JOINED.
DR EMBL; AB020844; BAA78577.1; JOINED.
DR EMBL; AB020845; BAA78577.1; JOINED.
DR EMBL; AB020846; BAA78577.1; JOINED.
DR EMBL; AB020847; BAA78577.1; JOINED.
DR EMBL; AB020848; BAA78577.1; JOINED.
DR EMBL; AB020849; BAA78577.1; JOINED.
DR EMBL; AB020850; BAA78577.1; JOINED.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 9.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00202; SR; 9.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; UNKNOWN_8.
DR PROSITE; PS00287; SRCR_2; 9.
DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
SQ SEQUENCE 1785 AA; 193991 MW; 38B2363F95226E80 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 1785;
Best Local Similarity 64.3%; Pred. No. 54;

```

Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYSSG 15  
DB 1134 PSSNCGGFLFYASG 1147

RESULT 5  
QY4V9  
ID QY4V9 PRELIMINARY; PRT: 1785 AA.  
AC QY4V9  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE DBMT1/8kb.1 protein precursor.  
GN DBMT1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
RC TISSUE=LUNG;  
RX MEDLINE=97434209; PubMed=9288095;  
RA Mollenhauer J., Wiemann S., Scheurle W., Korn B., Hayashi Y.,  
RA Willgenbus K.K., von Deimling A., Poustka A.;  
RT "DBMT1, a new member of the SRCR superfamily on chromosome 10q25.3-  
RT q26.1 is deleted in malignant brain tumours.";  
RL Nat. Genet. 17:32-39(1997).  
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
DR EMBL: AJ000342; CAA04019.1; -.  
DR InterPro: IPR000859; CUB\_domain.  
DR InterPro: IPR001507; Endoglin/CD105.  
DR InterPro: IPR001190; Srcr\_receptor.  
DR Pfam: PF00431; CUB; 2.  
DR Pfam: PF00530; SRCR; 9.  
DR Pfam: PF00100; zona\_pellucida; 1.  
DR PRINTS: PR00258; SPERACTRCPTR.  
DR SMART: SM00042; CUB; 2.  
DR SMART: SM00202; SR; 9.  
DR SMART: SM00241; ZP; 1.  
DR PROSITE: PS01180; CUB; 2.  
DR PROSITE: PS00420; SRCR\_1; UNKNOWN\_13.  
DR PROSITE: PS50287; SRCR\_2; 14.  
DR PROSITE: PS00682; ZP\_DOMAIN; UNKNOWN\_1.  
FT CHAIN 26 1785 DBMT1/8kb.1 PROTEIN.  
KW Signal.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 1785 DBMT1/8kb.1 PROTEIN.  
FT SEQUENCE 1785 AA; 193941 MW; 7B1F8D4/E4A82092 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 1785;  
Best Local Similarity 64.3%; Pred. No. 54;  
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYSSG 15  
DB 1134 PSSNCGGFLFYASG 1147

RESULT 6  
QY4V9  
ID QY4V9 PRELIMINARY; PRT: 2403 AA.  
AC QY4V9  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE DBMT1/8kb.1 protein.  
GN DBMT1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_TaxID=9606;  
[1]

RP SEQUENCE FROM N.A.  
RC TISSUE=LUNG;  
RX MEDLINE=20065089; PubMed=10597221;  
RA Mollenhauer J., Holmskov U., Wiemann S., Krebs I., Herberitz S.,  
RA Madsen J., Kioschis P., Coy J.F., Poustka A.;  
RT "The genomic structure of the DBMT1 gene: evidence for a region with  
RT susceptibility to genomic instability.";  
RL Oncogene 18:6233-6240(1999).  
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
DR EMBL: AJ243224; CAB63942.1; -.  
DR InterPro: IPR000859; CUB\_domain.  
DR InterPro: IPR001507; Endoglin/CD105.  
DR InterPro: IPR001190; Srcr\_receptor.  
DR Pfam: PF00431; CUB; 2.  
DR Pfam: PF00530; SRCR; 14.  
DR Pfam: PF00100; zona\_pellucida; 1.  
DR PRINTS: PR00258; SPERACTRCPTR.  
DR SMART: SM00042; CUB; 2.  
DR SMART: SM00202; SR; 14.  
DR SMART: SM00241; ZP; 1.  
DR PROSITE: PS01180; CUB; 2.  
DR PROSITE: PS00420; SRCR\_1; UNKNOWN\_13.  
DR PROSITE: PS50287; SRCR\_2; 14.  
DR PROSITE: PS00682; ZP\_DOMAIN; UNKNOWN\_1.  
FT CHAIN 26 2403 DBMT1/8kb.1 PROTEIN.  
SQ SEQUENCE 2403 AA; 259573 MW; 902622DE7399AD38 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 2403;  
Best Local Similarity 64.3%; Pred. No. 74;  
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYSSG 15  
DB 1752 PSSNCGGFLFYASG 1765

RESULT 7  
QY4V9  
ID QY4V9 PRELIMINARY; PRT: 2412 AA.  
AC QY4V9  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE DBMT1/8kb.2 protein precursor.  
GN DBMT1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
RC TISSUE=TRACHEA;  
RX MEDLINE=99415938; PubMed=10485905;  
RA Holmskov U., Mollenhauer J., Madsen J., Vitved L., Groenlund J.,  
RA Tornøe I., Kliem A., Reid K.B., Poustka A., Skjodt K.;  
RT "Cloning of gp-340, a putative opsonin receptor for lung surfactant  
RT protein D.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).  
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
DR EMBL: AJ243212; CAB56155.1; -.  
DR HSSP: P29392; ISFP.  
DR InterPro: IPR000859; CUB\_domain.  
DR InterPro: IPR001507; Endoglin/CD105.  
DR InterPro: IPR001190; Srcr\_receptor.  
DR Pfam: PF00431; CUB; 2.  
DR Pfam: PF00530; SRCR; 14.  
DR Pfam: PF00100; zona\_pellucida; 1.  
DR PRINTS: PR00258; SPERACTRCPTR.  
DR SMART: SM00042; CUB; 2.  
DR SMART: SM00202; SR; 14.  
DR SMART: SM00241; ZP; 1.  
DR PROSITE: PS01180; CUB; 2.  
DR PROSITE: PS00420; SRCR\_1; UNKNOWN\_13.

```

DR PROSITE: PS0287; SRCR_2; 14.
DR PROSITE: PS00682; ZP_DOMAIN; UNKNOWN_1.
KW SIGNAL. 1 25 POTENTIAL.
FT CHAIN 26 2412 DMBT1/8KB.2 PROTEIN.
FT CHAIN 26 2412 DMBT1/8KB.2 PROTEIN.
SQ SEQUENCE 2412 AA: 260568 MW: 35630CCBFFB18DD CRC64;

Query Match
Best Local Similarity 52.3%; Score 46; DB 4; Length 2412;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYYSSG 15
DB 1762 PSSNCGGFLFYASG 1775
II III I:II

RESULT 8
Q9UKJ4 PRELIMINARY; PRT; 2413 AA.
ID Q9UKJ4
AC Q9UKJ4
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Gp-340 variant protein.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=99415938; PubMed=10485905;
RA Holmskov U., Mollenhauer J., Madsen J., Vitved L., Gronlund J.,
RA Tornoe I., Klien A., Reid K.B., Poustka A., Skjoldt K.;
RT "Cloning of gp-340, a putative opsonin receptor for lung surfactant
RT protein D."
RL Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR InterPro: IPR000859; CUB_domain.
DR InterPro: IPR001507; Endoglin/CD105.
DR Pfam: PF00431; CUB; 2.
DR Pfam: PF00530; SRCR; 14.
DR PRINTS: PR00258; SPERACTRCPTR.
DR SMART: SM00042; CUB; 2.
DR SMART: SM00202; SR; 14.
DR SMART: SM00241; ZP; 1.
DR PROSITE: PS01180; CUB; 2.
DR PROSITE: PS00420; SRCR_1; UNKNOWN_13.
DR PROSITE: PS0287; SRCR_2; 14.
DR PROSITE: PS00682; ZP_DOMAIN; UNKNOWN_1.
SQ SEQUENCE 2413 AA: 260755 MW: 039544043CF4G3D4 CRC64;

Query Match
Best Local Similarity 52.3%; Score 46; DB 4; Length 2413;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYYSSG 15
DB 1762 PSSNCGGFLFYASG 1775
II III I:II

RESULT 9
Q96DU4 PRELIMINARY; PRT; 2413 AA.
ID Q96DU4
AC Q96DU4
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE DMBT1/8kb.2 protein precursor.
GN DMBT1.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Mollenhauer J.;
RT "Major subforms of DMBT1 are gastrointestinal mucins that display
RT extensive alternative splicing and differential protein targeting."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ297935; CAC44122.1; -.
DR InterPro: IPR000859; CUB_domain.
DR InterPro: IPR001507; Endoglin/CD105.
DR InterPro: IPR001190; Srcr_receptor.
DR Pfam: PF00431; CUB; 2.
DR Pfam: PF00530; SRCR; 14.
DR Pfam: PF00100; zona_pellucida; 1.
DR PROSITE: PS01180; CUB; 2; UNKNOWN_13.
DR PROSITE: PS00420; SRCR_1; UNKNOWN_13.
DR PROSITE: PS0287; SRCR_2; 14.
DR PROSITE: PS00682; ZP_DOMAIN; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 2413 DMBT1/8KB.2 PROTEIN
SQ SEQUENCE 2413 AA: 260600 MW: 99A449C5F4F60728 CRC64;

Query Match
Best Local Similarity 52.3%; Score 46; DB 4; Length 2413;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYYSSG 15
DB 1762 PSSNCGGFLFYASG 1775
II III I:II

RESULT 10
Q9UGM3 PRELIMINARY; PRT; 2426 AA.
ID Q9UGM3
AC Q9UGM3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE DMBT1 prototype precursor.
GN DMBT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20065089; PubMed=10597221;
RA Mollenhauer J., Holmskov U., Wiemann S., Krebs I., Herberitz S.,
RA Madsen J., Kioschis P., Coy J.F., Poustka A.;
RT "The genomic structure of the DMBT1 gene: evidence for a region with
RT susceptibility to genomic instability."
RL Oncogene 18:6233-6240(1999).
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL: AJ243211; CAB63941.1; -.
DR InterPro: IPR000859; CUB_domain.
DR InterPro: IPR001507; Endoglin/CD105.
DR InterPro: IPR001190; Srcr_receptor.
DR Pfam: PF00431; CUB; 2.
DR Pfam: PF00530; SRCR; 14.
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00258; SPERACTRCPTR.
DR SMART: SM00042; CUB; 2.
DR SMART: SM00202; SR; 14.
DR SMART: SM00241; ZP; 1.
DR PROSITE: PS01180; CUB; 2.
DR PROSITE: PS00420; SRCR_1; UNKNOWN_13.
DR PROSITE: PS0287; SRCR_2; 14.
DR PROSITE: PS00682; ZP_DOMAIN; UNKNOWN_1.

```

KW Signal.  
 FT SIGNAL 1 25 POTENTIAL.  
 ET CHAIN 26 2426 DMBP1 PROTOTYPE.  
 SQ SEQUENCE 2426 AA; 262052 MW; 5A58FBC076FB7247 CRC64;

Query Match 52.3%; Score 46; DB 4; Length 2426;  
 Best Local Similarity 64.3%; Pred. No. 75;  
 Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

OY 4 PSCNGC-LYSSG 15  
 || ||| | : | : ||  
 DB 1775 PSSNGGFLFVSG 1788

## RESULT 11

Q8TAE0 ID Q8TAE0 PRELIMINARY; PRT; 183 AA.  
 AC Q8TAE0  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Koyt binding protein 2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIMPH NODE;  
 RA Li R., Han H., Wang J.;  
 RT "KBP, a novel protein interacting with LIM protein Koyt.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF493784; AAM12863.1; -;  
 DR EMBL: AF493786; AAM12866.1; -;  
 SQ SEQUENCE 183 AA; 20207 MW; 872DD878B01C513B CRC64;

Query Match 51.1%; Score 45; DB 4; Length 183;  
 Best Local Similarity 66.7%; Pred. No. 7.6;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 YPSCNGGLYSS 14  
 | | | | | | | |  
 DB 95 YTSQCGKYSS 106

## RESULT 12

Q8TAC6 ID Q8TAC6 PRELIMINARY; PRT; 210 AA.  
 Q8TAC6  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Koyt binding protein 1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIMPH NODE;  
 RA Li R., Han H., Wang J.;  
 RT "KBP, a novel protein interacting with LIM protein Koyt.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF493783; AAM12862.1; -;  
 DR EMBL: AF493786; AAM12865.1; -;  
 SQ SEQUENCE 210 AA; 23114 MW; F00AA1F166C37967 CRC64;

Query Match 51.1%; Score 45; DB 4; Length 210;  
 Best Local Similarity 66.7%; Pred. No. 8.7;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 YPSCNGGLYSS 14  
 | | | | | | | |

DB 95 YTSQCGKYSS 106

## RESULT 13

Q9CU50 ID Q9CU50 PRELIMINARY; PRT; 279 AA.  
 AC Q9CU50  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE 6330404ELLIRIK protein (Fragment).  
 GN 6330404ELLIRIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=MEDULLA OBLONGATA;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,  
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK018113; BAB31076.1; -;  
 DR HSSP: P02671; LEZD.  
 DR MGD: MGI:1917976; 6330404ELLIRIK.  
 DR InterPro: IPR002181; Fibrinogen\_C.  
 DR Pfam: PF00147; Fibrinogen\_C; 1.  
 DR SMART: SM00186; FBG; 1.  
 DR PROSITE: PS00514; FIBRIN\_AG\_C\_DOMAIN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 279 279  
 SQ SEQUENCE 279 AA; 31507 MW; D37B69F61F69C9A5 CRC64;

Query Match 50.6%; Score 44.5; DB 11; Length 279;  
 Best Local Similarity 60.0%; Pred. No. 14;  
 Matches 9; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 DRYSCNGGLYSS 15  
 | | | | | | | |  
 DB 219 DSY-SCNCLYHRCG 232

## RESULT 14

Q8R0Z6 ID Q8R0Z6 PRELIMINARY; PRT; 457 AA.  
 AC Q8R0Z6  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Similar to angiopoietin-related protein 5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

|                            |   |
|----------------------------|---|
| RP                         | SEQUENCE FROM N.A.  |
| RC                         | TISSUE=LIVER;   |
| RA                         | Strausberg R.;  |
| RL                         | Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.          |
| RQ                         | EMBL; BC025904; AA#25904.1; "                                     |
| SQ                         | SEQUENCE 457 AA; 51095 MW; B7C4289E3FEC6C3E CRC64;                |
| <br>                       |   |
| Query Match                | 50.6%; Score 44.5; DB 11; Length 457;                             |
| Best Local Similarity      | 60.0%; Pred. No. 24;  |
| Matches 9; Conservative 1; | Mismatches 4; Indels 1; Gaps 1;                                   |
| <br>                       |   |
| QY                         | 1 DRYPSGNCGLYYSSG 15<br>I I I I I I I I I                         |
| DB                         | 391 DSY-SGCALYHRGG 404  |
| <br>                       |   |
| RESULT 15                  |   |
| Q9GJH7                     | PRELIMINARY; PRT; 217 AA.   |
| ID                         | Q9GJH7  |
| NC                         | Q9GJH7;   |
| CC                         | 01-MAR-2001 (TrEMBLrel. 16, Created)                              |
| CD                         | 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)                 |
| DT                         | 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)               |
| DE                         | MHC class II beta-chain (fragment).                               |
| DS                         | SATR-DAB.   |
| GN                         | Salmo trutta (Brown trout).                                       |
| OS                         | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC                         | Aktinopterygii; Neopterygii; Teleostei; Euteleostei;              |
| OX                         | Protactinopterygii; Salmoniformes; Salmonidae; Salmo.             |
| OX                         | NCB1_TaxID=8032;  |
| OR                         | [1]   |
| RN                         | SEQUENCE FROM N.A.  |
| RC                         | TISSUE=SPLREEN;   |
| RX                         | MEDLINE-21142794; PubMed-11207285;                                |
| RA                         | Shum B.P., Guethlein L., Flodin L.R., Adkison M.A., Hedrick R.P., |
| RL                         | Nehring R.B., Stet R.J.M., Secombes C., Parham P.;                |
| RM                         | "Modes of Salmonid MHC Class I and II Evolution Differ from the   |
| RT                         | Primate Paradigm.";   |
| RU                         | J. Immunol. 166:3297-3308(2001).                                  |
| DR                         | EMBL; AF256339; AAC02545.1; "-;                                   |
| DR                         | HSP; P01B88; IBMG.  |
| DR                         | InterPro: IPR003597; Ig-cl.                                       |
| DR                         | InterPro: IPR003006; Ig_MHC.                                      |
| DR                         | InterPro: IPR000353; MHC_II_beta.                                 |
| DR                         | Pfam: PF00047; ig; 1.   |
| DR                         | Pfam: PF00969; MHC_II_beta; 1.                                    |
| DR                         | ProDom: PD000328; MHC_II_beta; 1.                                 |
| DR                         | SMART; SM00407; IGcl; 1.  |
| DR                         | Glycoprotein; MHC II; Transmembrane.                              |
| SV                         | NON_TER 1   |
| FT                         | NON_TER 217 217   |
| SQ                         | SEQUENCE 217 AA; 24407 MW; 7DFC6B63EE619F8C CRC64;                |
| <br>                       |   |
| Query Match                | 50.0%; Score 44; DB 7; Length 217;                                |
| Best Local Similarity      | 50.0%; Pred. No. 13;  |
| Matches 7; Conservative 3; | Mismatches 4; Indels 0; Gaps 0;                                   |
| <br>                       |   |
| QY                         | 1 DRYPSGNCGLYYSS 14<br>: : : : : : : : :                          |
| DB                         | 69 ERYCKGNAAIYYSA 82  |

Search completed: December 2, 2002, 06:53:26  
Job time : 7.09692 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:38:29 ; Search time 2.11454 Seconds  
(without alignments)  
945.247 Million cell updates/sec

Title: US-09-902-563-18  
Perfect score: 88  
Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
al number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Query Match | Score | Length | DB ID | Description                 |
|------------|-------------|-------|--------|-------|-----------------------------|
| 1          | 88          | 100.0 | 15     | 20    | AAW88237 Human prothrombina |
| 2          | 88          | 100.0 | 432    | 20    | AAW88236 Mouse prothrombina |
| 3          | 88          | 100.0 | 439    | 20    | AAW88235 Human prothrombina |
| 4          | 50          | 56.8  | 53     | 23    | ABP34223 Human angiotensin  |
| 5          | 47          | 53.4  | 93     | 22    | AAE03347 Human gene 1 encod |
| 6          | 46          | 52.3  | 102    | 22    | ABG25291 Novel human diagno |
| 7          | 46          | 52.3  | 666    | 19    | AAW64590 Human SRCR protein |
| 8          | 46          | 52.3  | 1785   | 19    | AAW64591 Human SRCR protein |
| 9          | 46          | 52.3  | 3460   | 23    | ABB05007 Human reelin prote |
| 10         | 46          | 52.3  | 3461   | 23    | ABB05008 Mouse reelin prote |

|    |      |      |      |    |          |                     |
|----|------|------|------|----|----------|---------------------|
| 11 | 46   | 52.3 | 3461 | 23 | ABB57065 | Mouse ischaemic co  |
| 12 | 46   | 52.3 | 3470 | 22 | ABG25297 | Novel human diagno  |
| 13 | 45   | 51.1 | 123  | 22 | AAU80079 | Apoptin-associatin  |
| 14 | 45   | 51.1 | 126  | 22 | AAU80078 | Apoptin-associatin  |
| 15 | 45   | 51.1 | 126  | 23 | AAO14806 | Human apoptin-asso  |
| 16 | 45   | 51.1 | 158  | 21 | AA42752  | Human ORFX ORF2516  |
| 17 | 45   | 51.1 | 210  | 22 | AAU80085 | Apoptin-associatin  |
| 18 | 45   | 51.1 | 210  | 22 | AAU80085 | Human polyptide,    |
| 19 | 44   | 50.0 | 292  | 22 | ABB70950 | Drosophila melanog  |
| 20 | 43   | 48.9 | 460  | 20 | AAV34589 | Chlamydia pneumoni  |
| 21 | 42.5 | 48.3 | 193  | 18 | AAW53378 | H. pylori ORF 07cp  |
| 22 | 42.5 | 48.3 | 193  | 23 | AAW50313 | Helicobacter pylori |
| 23 | 42.5 | 48.3 | 493  | 18 | AAW55517 | H. pylori ORF 07ce  |
| 24 | 42.5 | 48.3 | 493  | 23 | AAW50311 | Helicobacter pylori |
| 25 | 42.5 | 48.3 | 696  | 18 | AAW55698 | H. pylori ORF 14gp  |
| 26 | 42.5 | 48.3 | 696  | 20 | AAV17207 | H. pylori outer me  |
| 27 | 42.5 | 48.3 | 696  | 23 | AAW50309 | Helicobacter pylori |
| 28 | 42.5 | 48.3 | 696  | 23 | AAW50310 | Helicobacter pylori |
| 29 | 41.5 | 47.2 | 279  | 22 | AAV72625 | Human angiotensin   |
| 30 | 41.5 | 47.2 | 342  | 22 | AAV72626 | Human angiotensin   |
| 31 | 41.5 | 47.2 | 470  | 20 | AAI05398 | Human TIE ligand N  |
| 32 | 41.5 | 47.2 | 470  | 21 | AAE24389 | Human PRO178 prote  |
| 33 | 41.5 | 47.2 | 470  | 22 | AAE51330 | Human NEW angioten  |
| 34 | 41.5 | 47.2 | 470  | 22 | AAE53066 | Human angiotensin   |
| 35 | 41.5 | 47.2 | 470  | 23 | ABE95424 | Human angiotensin   |
| 36 | 41.5 | 47.2 | 470  | 23 | AAE19827 | Human TIE ligand N  |
| 37 | 41.5 | 47.2 | 470  | 23 | ABE84818 | Human PRO178 prote  |
| 38 | 41   | 46.6 | 39   | 17 | AAW05340 | Calisoga spider ve  |
| 39 | 41   | 46.6 | 39   | 17 | AAW05341 | Calisoga spider ve  |
| 40 | 41   | 46.6 | 80   | 17 | AAW05343 | Calisoga spider ve  |
| 41 | 41   | 46.6 | 93   | 22 | AAW90273 | Human immune/haema  |
| 42 | 41   | 46.6 | 213  | 22 | ABG27841 | Novel human diagno  |
| 43 | 41   | 46.6 | 239  | 21 | AAE41169 | Human ORFX ORF933   |
| 44 | 41   | 46.6 | 637  | 22 | ABE60216 | Drosophila melanog  |
| 45 | 41   | 46.6 | 645  | 20 | AAV33681 | D. melanogaster Oc  |

## ALIGNMENTS

RESULT 1  
AAW88237  
ID AAW88237 standard; Peptide; 15 AA.  
XX AAW88237;  
AC AAW88237;  
XX 15-MAR-1999 (first entry)  
DT Human prothrombinase Fgl2 epitope.  
XX  
DE Prothrombinase: hfgl2; Fgl2; human; immune coagulation; antibody;  
KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;  
KW gastrointestinal disease; foetal loss; therapy; vaccine; epitope.  
XX  
OS Homo sapiens.  
XX  
XX WO9851335-A1.  
PN  
XX  
PD 19-NOV-1998.  
XX  
PF 15-MAY-1998; 98WO-CA00475.  
XX  
PR 10-OCT-1997; 97US-0061684.  
PR 15-MAY-1997; 97US-0046537.  
XX (LEVY/) LEVY G.  
XX  
XX Levy G;  
XX WPI; 1999-059687/05.  
XX  
XX Modulating immune coagulation - by using Fgl2 antibodies and  
PT compounds, used to treat conditions including graft rejection and

PT foetal loss  
 XX Claim 4; Page 72; 105pp; English.  
 XX This peptide corresponds to amino acid residues 364-378 of human  
 CC prothrombinase Fg12 (see AAW8235). A claimed method of preventing  
 CC or treating a condition requiring a reduction in immune  
 CC coagulation comprises administering an inhibitor of Fg12. The  
 CC inhibitor is preferably an antibody that binds to the Fg12 epitope.  
 CC The condition to be treated in graft rejection of foetal loss  
 CC (claimed).  
 XX  
 XX Sequence 15 AA;  
 SQ Query Match 100.0%; Score 88; DB 20; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 DRYPSGNCGLYSSG 15  
 ID AAW8235  
 AC AAW8236 standard; Protein; 432 AA.  
 XX AAW8236;  
 DT 15-MAR-1999 (first entry)  
 DE Mouse prothrombinase Fg12 protein.  
 XX  
 KW Prothrombinase; Fg12; mouse; immune coagulation; antibody;  
 KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;  
 KW gastrointestinal disease; foetal loss; therapy; vaccine.  
 XX  
 OS Mus sp.  
 PH Key Location/Qualifiers  
 FT Modified-site 172..174  
 FT /label= Asn is N-glycosylated  
 FT Modified-site 228..231  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 256..259  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 323..325  
 FT /note= "Asn is N-glycosylated"  
 FT Domain 213..439  
 FT /note= "fibrinogen related domain"  
 XX  
 PN WO9851335-A1.  
 XX  
 PD 19-NOV-1998.  
 XX  
 PF 15-MAY-1998; 98WO-CA00475.  
 XX  
 PR 10-OCT-1997; 97US-0061684.  
 PR 15-MAY-1997; 97US-0046537.  
 XX  
 PA (LEVY/) LEVY G.  
 XX  
 PI Levy G;  
 XX  
 DR WPI: 1999-059687/05.  
 DR N-PSDB; AAW84140.  
 XX  
 PT Modulating immune coagulation - by using Fg12 antibodies and  
 PT compounds, used to treat conditions including graft rejection and  
 PT foetal loss  
 XX  
 PS Claim 8; Page 70-71; 105pp; English.  
 XX

CC This is the amino acid sequence of mouse prothrombinase Fg12, as  
 CC predicted from fg12 DNA (see AAW84140). Fg12 is a 70 kDa  
 CC transmembrane serine protease that has immune procoagulant activity.  
 CC The human Fg12 amino acid sequence is given in AAW8236. The  
 CC invention provides a method for inhibiting immune coagulation by  
 CC inhibiting the activity or expression of Fg12. The method can be  
 CC used in vivo to treat a condition which requires a reduction in  
 CC immune coagulation such as bacterial and viral infections, cancer,  
 CC glomerulonephritis, a number of gastrointestinal diseases,  
 CC allograft and xenograft rejection and foetal loss. An Fg12-specific  
 CC antibody, an Fg12 antisense oligonucleotide, or a substance that  
 CC affects prothrombinase activity of a Fg12 protein may be used to  
 CC treat a condition requiring a reduction in procoagulant activity.  
 CC A vaccine containing an Fg12 protein or peptide is used for  
 CC prevention of graft rejection or foetal loss (claimed).  
 XX  
 XX Sequence 432 AA;  
 SQ Query Match 100.0%; Score 88; DB 20; Length 432;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 DRYPSGNCGLYSSG 15  
 ID AAW8235  
 AC AAW8235 standard; Protein; 439 AA.  
 XX AAW8235;  
 DT 15-MAR-1999 (first entry)  
 DE Human prothrombinase Fg12 protein.  
 XX  
 KW Prothrombinase; hfg12; human; immune coagulation; antibody;  
 KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;  
 KW gastrointestinal disease; foetal loss; therapy; vaccine.  
 XX  
 OS Homo sapiens.  
 PH Key Location/Qualifiers  
 FT Modified-site 179..183  
 FT /label= Asn is N-glycosylated  
 FT Modified-site 235..238  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 262..265  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 336..337  
 FT /note= "Asn is N-glycosylated"  
 FT Domain 213..439  
 FT /note= "fibrinogen related domain"  
 FT Peptide 364..378  
 FT /note= "epitope (Claim 4)"  
 XX  
 PN WO9851335-A1.  
 XX  
 PD 19-NOV-1998.  
 XX  
 PF 15-MAY-1998; 98WO-CA00475.  
 XX  
 PR 10-OCT-1997; 97US-0061684.  
 PR 15-MAY-1997; 97US-0046537.  
 XX  
 PA (LEVY/) LEVY G.  
 XX  
 PI Levy G;  
 XX  
 DR WPI: 1999-059687/05.  
 DR N-PSDB; AAW84139.  
 XX

PT Modulating immune coagulation - by using Fgl2 antibodies and  
 PT compounds, used to treat conditions including graft rejection and  
 PT foetal loss

PS Claim 8: Page 66-67; 105pp; English.

XX This is the amino acid sequence of human prothrombinase Fgl2, as  
 CC predicted from hfgl2 DNA (see AAV84139). Fgl2 is a 70 kDa  
 CC transmembrane serine protease that has immune procoagulant activity.  
 CC The invention provides a method for inhibiting immune coagulation by  
 CC inhibiting the activity or expression of Fgl2. The method can be  
 CC used in vivo to treat a condition which requires a reduction in  
 CC immune coagulation such as bacterial and viral infections, cancer,  
 CC glomerulonephritis, a number of gastrointestinal diseases,  
 CC allograft and xenograft rejection and foetal loss. An Fgl2-specific  
 CC antibody, an Fgl2 antisense oligonucleotide, or a substance that  
 CC affects prothrombinase activity of a Fgl2 protein may be used to  
 CC treat a condition requiring a reduction in procoagulant activity.  
 CC A vaccine containing an Fgl2 protein or peptide is used for  
 CC prevention of graft rejection or foetal loss (claimed).

Sequence 439 AA;

Query Match 100.0%; Score 88; DB 20; Length 439;  
 Best Local Similarity 100.0%; Pred. No. 3e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSNGCLYYSSG 15  
 DB 364 DRYPSNGCLYYSSG 378  
 |||||

RESULT 4

ABP34223  
 ID ABP34223 standard; Protein: 53 AA.

AC ABP34223;

XX 08-JUL-2002 (first entry)

XX Human angiotensin-like ORF3196 protein, SEQ ID NO:6392.

XX Human: ORF; open reading frame: ORFX; drug screening; diagnosis;  
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;  
 KW immune modulation; haematopoiesis regulation; tissue growth;  
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;  
 KW behaviour; cancer; proliferative disorder; neurological disorder;  
 KW cardiovascular disease; immune system disorder; organ transplantation;  
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;  
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;  
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;  
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX Homo sapiens.

XX WO200190366-A2.

XX 29-NOV-2001.

PF 24-MAY-2001: 2001WO-US17076.

PR 24-MAY-2000: 2000US-206690P.

PA (CURA-) CURAGEN CORP.

PI Leach MD, Shimkets RA;

XX WPI: 2002-106200/14.

DR N-PSDB: ABN78249.

PT Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and disorders related to organ  
 PT transplantation

PS Claim 10: Page 1846; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins  
 CC designated ORF (Open reading frame) 1-4534, and sequences ABN75054-  
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses  
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX  
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
 CC polypeptides, methods of screening for modulators of ORFX expression or  
 CC activity, and methods of screening individuals for a predisposition to an  
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
 CC range of biological activities, such as cytokine, cell proliferation,  
 CC cell differentiation, immune modulation, haematopoiesis regulation,  
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
 CC and antiinfective activity, and may also be involved in the determination  
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
 CC nucleic acids and antibodies may be used in the treatment of cancers,  
 CC other proliferative disorders such as psoriasis and benign tumours,  
 CC neurological disorders such as epilepsy and Alzheimer's disease,  
 CC cardiovascular diseases, immune system disorders, disorders related to  
 CC organ transplantation, disorders of tissue growth and regeneration,  
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester  
 CC storage disease, and infectious diseases caused by viral, bacterial,  
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
 CC source of primers and probes, in the detection of ORFX genomic sequences  
 CC or transcripts, in the identification and cloning of homologous  
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
 CC nucleic acids may additionally be used to produce transgenic animals  
 CC which may be useful for studying the function and/or activity of ORFX  
 CC protein, and in drug screening. The ORFX proteins may also be used as  
 CC immunogens to generate specific antibodies, which are useful in the  
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX Sequence 53 AA;

Query Match 56.8%; Score 50; DB 23; Length 53;  
 Best Local Similarity 66.7%; Pred. No. 2;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PSNGCLYYSSG 15  
 DB 20 PSNGCLYHGG 31  
 |||||

RESULT 5  
 AAE03347  
 ID AAE03347 standard; peptide; 93 AA.

XX AC AAE03347;

XX 10-AUG-2001 (first entry)

XX Human gene 1 encoded secreted protein fragment, SEQ ID NO:122.

XX Human; secreted protein; proliferative disorder; cancer; tumour;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
 KW cardiovascular disorder; angiotensin disorder; kidney disorder;  
 KW gastrointestinal disorder; pregnancy-related disorder;  
 KW endocrine disorder; infection; wound healing; vulnery;  
 KW cell culture; chemotaxis; food additive; gene therapy;

KW binding partner identification.  
 XX Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 7  
 FT /label= Unknown  
 FT /note= "X equals stop translation"  
 XX WO200134800-A1.  
 XX 17-MAY-2001.  
 XX  
 XX 08-NOV-2000; 2000WO-US30674.  
 XX  
 XX 12-NOV-1999; 99US-0164750.  
 XX 30-JUN-2000; 2000US-0215128.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Komatsoulis CA, Ebner R, Fiscella M, Wei P;  
 XX WPI; 2001-329085/34.  
 XX  
 XX New nucleic acid molecules encoding human secreted proteins, used in  
 XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
 XX Parkinson's diseases and cancers  
 XX  
 XX Disclosure; Page 501; 530pp; English.  
 XX  
 XX AAD07705-AAD07759 represent cDNAs corresponding to 19 human secreted  
 XX protein genes, and AAE03292-AAE03346 represent the proteins they encode.  
 XX AAE03347-AAE03375 represent human secreted protein fragments or variants.  
 XX The genes and their secreted proteins are useful for preventing,  
 XX treating or ameliorating medical conditions, e.g., by protein or gene  
 XX therapy. Pathological conditions can be diagnosed by determining the  
 XX amount of the new protein in a sample or by determining the presence of  
 XX mutations in the new genes. Specific uses are described for each of the  
 XX 19 genes, based on the tissues in which they are most highly expressed,  
 XX and include developing products for the diagnosis or treatment of  
 XX proliferative disorders, cancer, tumours, foetal and developmental  
 XX abnormalities, haematopoietic disorders, diseases of the immune system,  
 XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,  
 XX allergies, neurological disorders (e.g., Alzheimer's disease,  
 XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,  
 XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
 XX cardiovascular disorders, angiogenic disorders, kidney disorders,  
 XX gastrointestinal disorders, pregnancy-related disorders, endocrine  
 XX disorders, and infections. The proteins can also be used to aid wound  
 XX healing and epithelial cell proliferation, to prevent skin aging due to  
 XX sunburn, to maintain organs before transplantation, for supporting cell  
 XX culture of primary tissues, to regenerate tissues, to identify their  
 XX cognate ligands or binding partners, and in chemotaxis, and can be used  
 XX as a food additive or preservative to modify storage properties.  
 XX Antibodies specific for a protein of the invention can be used in  
 XX alleviating symptoms associated with the disorders mentioned above, and  
 XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
 XX immunosorbent assay (ELISA). The present sequence represents a human  
 XX secreted protein fragment referred to in the disclosure of the invention.  
 XX  
 XX SQ Sequence 93 AA;  
 XX  
 XX Query Match 53.4%; Score 47; DB 22; Length 93;  
 XX Best Local Similarity 66.7%; Pred. No. 9.9;  
 XX Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 4 PSNGCGLYVSSG 15  
 DB 32 PSNGCALYQRRG 43  
 RESULT 6  
 ABG25291

ID ABG25291 standard; Protein; 102 AA.  
 XX  
 AC ABG25291;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #25282.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 XX 30-MAR-2001; 2001WO-US08631.  
 XX  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Drmanac RT, Liu C, Tang YT;  
 XX  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS89478.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 XX diagnostics, forensics, gene mapping, identification of mutations  
 XX responsible for genetic disorders or other traits and to assess  
 XX biodiversity  
 XX  
 PS Claim 20; SEQ ID No 55650; 103pp; English.  
 XX  
 XX The invention relates to isolated polynucleotide (I) and  
 XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 XX and gene mapping, and in recombinant production of (II). The  
 XX polynucleotides are also used in diagnostics as expressed sequence tags  
 XX for identifying expressed genes. (I) is useful in gene therapy techniques  
 XX to restore normal activity of (II) or to treat disease states involving  
 XX quantitating a polypeptide in tissue, as molecular weight markers and as  
 XX a food supplement. (II) and its binding partners are useful in medical  
 XX imaging of sites expressing (II). (I) and (II) are useful for treating  
 XX disorders involving aberrant protein expression or biological activity.  
 XX The polypeptide and polynucleotide sequences have applications in  
 XX diagnostics, forensics, gene mapping, identification of mutations  
 XX responsible for genetic disorders or other traits to assess biodiversity  
 XX and to produce other types of data and products dependent on DNA and  
 XX amino acid sequences. ABG0010-ABG0377 represent novel human  
 XX diagnostic amino acid sequences of the invention.  
 XX Note: The sequence data for this patent did not appear in the printed  
 XX specification, but was obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX SQ Sequence 102 AA;  
 XX  
 XX Query Match 52.3%; Score 46; DB 22; Length 102;  
 XX Best Local Similarity 71.4%; Pred. No. 15;  
 XX Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;  
 QY 3 YPSGNCGLY--YSS 14  
 DB 86 YPGGNIGLYCPYSS 99  
 RESULT 7  
 AAW64590  
 ID AAW64590 standard; Protein; 666 AA.  
 XX

AC AAW64590;  
 XX  
 DP 23-OCT-1998 (first entry)  
 XX  
 DE Human SRCR protein fragment.  
 XX  
 KW Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;  
 KW nervous system; medullo-blastoma; glioma; breast; detection;  
 KW autoantibody.  
 XX  
 OS Homo sapiens.  
 XX  
 EH Key Location/Qualifiers  
 FT Misc-difference 3  
 FT /label= unknown  
 XX  
 PN WO9830687-A2.  
 XX  
 PD 16-JUL-1998.  
 XX  
 PY 09-JAN-1998; 98WO-DE000096.  
 PR 18-JUL-1997; 97DE-1030997.  
 PR 09-JAN-1997; 97DE-1000519.  
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 XX  
 PI Mollenhauer J, Poustka A;  
 XX  
 DR WPI; 1998-399136/34.  
 DR N-PSDB; AAV49651.  
 XX  
 PT Proteins containing scavenger receptor, cysteine rich domain -  
 PT useful for diagnosis and treatment of tumours  
 XX  
 PS Claim 1; Fig 1; 54pp; German.  
 XX  
 CC This sequence represents a fragment of a human protein which contains a  
 CC SRCR (scavenger receptor, cysteine-rich) domain. The gene and encoded  
 CC protein can be used to diagnose or treat tumours, particularly of the  
 CC nervous system (medullo-blastoma or glioma) or breast. The DNA sequence  
 CC and probes derived from it, are used to identify genes that express  
 CC SRCR-domain containing proteins, to determine the form in which these  
 CC proteins exist and to assess the significance of individual forms on  
 CC cellular properties. The protein can be used to detect the presence of  
 CC autoantibodies, and Ab which regulate its expression.  
 XX  
 SO Sequence 666 AA;  
 XX  
 Query Match 52.3%; Score 46; DB 19; Length 666;  
 Best Local Similarity 64.3%; Pred. No. 1e+02;  
 Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;  
 XX  
 QY 4 PSNGCG--LYSSG 15  
 DB 183 PSSNGGFLFYASG 196  
 XX  
 RESULT 8  
 AAW64591  
 ID AAW64591 standard; Protein; 1785 AA.  
 XX  
 AC AAW64591;  
 XX  
 DP 23-OCT-1998 (first entry)  
 XX  
 DE Human SRCR protein.  
 XX  
 KW Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;  
 KW nervous system; medullo-blastoma; glioma; breast; detection;  
 KW autoantibody; ss.  
 XX  
 OS Homo sapiens.

XX WO9830687-A2.  
 PN  
 XX 16-JUL-1998.  
 PD  
 XX  
 PY 09-JAN-1998; 98WO-DE000096.  
 PR 18-JUL-1997; 97DE-1030997.  
 PR 09-JAN-1997; 97DE-1000519.  
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 XX  
 PI Mollenhauer J, Poustka A;  
 XX  
 DR WPI; 1998-399136/34.  
 DR N-PSDB; AAV49652.  
 XX  
 PT Proteins containing scavenger receptor, cysteine rich domain -  
 PT useful for diagnosis and treatment of tumours  
 XX  
 PS Claim 2; Fig 2; 54pp; German.  
 XX  
 CC This sequence represents a human protein which contains a SRCR (scavenger  
 CC receptor, cysteine-rich) domain. The gene and encoded protein can be used  
 CC to diagnose or treat tumours, particularly of the nervous system  
 CC (medullo-blastoma or glioma) or breast. The DNA sequence and probes  
 CC derived from it, are used to identify genes that express SRCR-domain  
 CC containing proteins, to determine the form in which these proteins exist  
 CC and to assess the significance of individual forms on cellular  
 CC properties. The protein can be used to detect the presence of  
 CC autoantibodies and antibodies which regulate its expression.  
 XX  
 SO Sequence 1785 AA;  
 XX  
 Query Match 52.3%; Score 46; DB 19; Length 1785;  
 Best Local Similarity 64.3%; Pred. No. 2.7e+02;  
 Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;  
 XX  
 QY 4 PSNGCG--LYSSG 15  
 DB 1134 PSSNGGFLFYASG 1147  
 XX  
 RESULT 9  
 ABB05007  
 ID ABB05007 standard; Protein; 3460 AA.  
 XX  
 AC ABB05007;  
 XX  
 DP 21-MAR-2002 (first entry)  
 DT  
 XX  
 DE Human reelin protein SEQ ID NO:1.  
 XX  
 KW Human; reelin; low density lipoprotein receptor; LDLR; neuroprotective;  
 KW extracellular glycoprotein; nootropic; antilipemic; Alzheimer's disease;  
 KW neurodegenerative disorder; neuronal regeneration; cognitive function;  
 KW lipid metabolism disease; memory; developmental disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6323177-B1.  
 XX  
 PD 27-NOV-2001.  
 XX  
 PY 16-JUN-1999; 99US-0334220.  
 PR 16-JUN-1999; 99US-0334220.  
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Curran T, D'Arcangelo G;  
 XX  
 DR WPI; 2002-096596/13.

DR N-PSDB; ABA92604.

XX Novel composition useful for screening compounds that modulate Reelin  
PT binding to low density lipoprotein receptor, comprising an isolated  
PT Reelin polypeptide and low density lipoprotein receptor -  
XX

PS Claim 16; Column 31-48; 45pp; English.

XX The present invention describes a composition (I) comprising an  
CC isolated reelin protein (II) bound to an isolated low density lipoprotein  
CC receptor (LDLR) (III). (II) is an extracellular glycoprotein of  
CC approximately 385 kDa containing a small region of similarity with  
CC F-spondin at the N terminus, a stretch of positively charged amino  
CC acids at the C terminus, and a series of eight internal repeats of  
CC 350-390 amino acids, each repeat containing two related sub-domains  
CC that flank a pattern of conserved cysteine residues known as an  
CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,  
CC nontropic and antilipemic activities, and can be used as a modulator  
CC of reelin-LDLR interaction. (I) is useful in screen for compounds that  
CC modulate reelin binding to an LDLR, in an assay system, where the assay  
CC system comprises a microplate array and an automated robotic  
CC microprocessor controlled system for adding and removing reagents to  
CC the microplate array. The compounds identified by the above screening  
CC method are useful as therapeutic agents to provide or alleviate a  
CC diverse spectrum of diseases including neurodegenerative disorders such  
CC as Alzheimer's disease, to facilitate neuronal regeneration after  
CC injury, to prevent or alleviate lipid metabolism diseases, to enhance  
CC cognitive functions and memory or to ameliorate other developmental  
CC disorders. The present sequence represents human reelin, which is used  
CC in the exemplification of the present invention.

XX Sequence 3460 AA;

Query Match 52.3%; Score 46; DB 23; Length 3460;  
Best Local Similarity 71.4%; Pred. No. 5.3e+02;  
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPGNCGLY--YSS 14  
II II III III  
DB 1973 YPGNIGLYCPYSS 1986

RESULT 10

ABB05008  
ID ABB05008 standard; Protein; 3461 AA.

XX

AC ABB05008;

XX 21-MAR-2002 (first entry)

DE Mouse reelin protein SEQ ID NO:2.

XX Mouse; reelin; low density lipoprotein receptor; LDLR; neuroprotective;  
KW extracellular glycoprotein; nontropic; antilipemic; Alzheimer's disease;  
KW neurodegenerative disorder; neuronal regeneration; cognitive function;  
KW lipid metabolism disease; memory; developmental disorder.

XX Mus musculus.

XX US6323177-B1.

XX 27-NOV-2001.

XX 16-JUN-1999; 99US-0334220.

XX 16-JUN-1999; 99US-0334220.

XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Cuaran T, D'Arcangelo G;

XX WPI; 2002-096596/13.

DR N-PSDB; ABA92603.

XX

PT Novel composition useful for screening compounds that modulate Reelin  
PT binding to low density lipoprotein receptor, comprising an isolated  
PT Reelin polypeptide and low density lipoprotein receptor -  
XX

PS Claim 16; Column 47-64; 45pp; English.

XX The present invention describes a composition (I) comprising an  
CC isolated reelin protein (II) bound to an isolated low density lipoprotein  
CC receptor (LDLR) (III). (II) is an extracellular glycoprotein of  
CC approximately 385 kDa containing a small region of similarity with  
CC F-spondin at the N terminus, a stretch of positively charged amino  
CC acids at the C terminus, and a series of eight internal repeats of  
CC 350-390 amino acids, each repeat containing two related sub-domains  
CC that flank a pattern of conserved cysteine residues known as an  
CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,  
CC nontropic and antilipemic activities, and can be used as a modulator  
CC of reelin-LDLR interaction. (I) is useful in screen for compounds that  
CC modulate reelin binding to an LDLR, in an assay system, where the assay  
CC system comprises a microplate array and an automated robotic  
CC microprocessor controlled system for adding and removing reagents to  
CC the microplate array. The compounds identified by the above screening  
CC method are useful as therapeutic agents to provide or alleviate a  
CC diverse spectrum of diseases including neurodegenerative disorders such  
CC as Alzheimer's disease, to facilitate neuronal regeneration after  
CC injury, to prevent or alleviate lipid metabolism diseases, to enhance  
CC cognitive functions and memory or to ameliorate other developmental  
CC disorders. The present sequence represents mouse (Mus musculus) reelin,  
CC which is used in the exemplification of the present invention.

XX Sequence 3461 AA;

Query Match 52.3%; Score 46; DB 23; Length 3461;  
Best Local Similarity 71.4%; Pred. No. 5.3e+02;  
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPGNCGLY--YSS 14  
II II III III  
DB 1974 YPGNIGLYCPYSS 1987

RESULT 11

ABB57065

ID ABB57065 standard; Protein; 3461 AA.

XX

AC ABB57065;

XX 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related protein sequence SEQ ID NO:129.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.

XX Mus musculus.

XX WC2001188188-A2.

XX 22-NOV-2001.

XX 18-MAY-2001; 2001WO-JP04192.

XX 18-MAY-2000; 2000JP-0145977.

XX (UVNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX WPI; 2002-034733/04.

DR N-PSDB; ABI99284.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
PT expression levels of particular genes defined in the specification or

PT by determining the expression profile of a gene group comprising these  
 PT genes -  
 XX

PS Claim 2; Page 385-400; 2690pp; English.

XX The present invention describes a method for examining ischaemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (I) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (I). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (ABI99202 to ABI9912, encoding  
 CC the protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The  
 CC expression levels or expression profiles produced by these genes are  
 CC used as an indicator when screening for ischaemic condition-improving  
 CC drugs or therapeutics for ischaemic diseases. ABI9913 and ABI9914  
 CC represent PCR primers for a mouse ischaemic condition related sequence,  
 CC which are used in the exemplification of the present invention.  
 XX

XX Sequence 3461 AA;

Query Match 52.3%; Score 46; DB 23; Length 3461;  
 Best Local Similarity 71.4%; Pred. No. 5.3e+02;  
 Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNGCLY--YSS 14  
 ||||| |||||  
 Db 1974 YPGNGIGLYCPYSS 1987

RESULT 12

ABG25297  
 ID ABG25297 standard; Protein; 3470 AA.

XX

AC ABG25297;

XX

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #25288.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

OS

XX WO200175067-A2.

PN

XX 11-OCT-2001.

XX

XX 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

XX

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI: 2001-639362/73.

XX

DR N-PSDB; AAS89484.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations,

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

PS Claim 20; SEQ ID NO 55656; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG030377 represent novel human  
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 3470 AA;

Query Match 52.3%; Score 46; DB 22; Length 3470;  
 Best Local Similarity 71.4%; Pred. No. 5.3e+02;  
 Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNGCLY--YSS 14  
 ||||| |||||  
 Db 1983 YPGNGIGLYCPYSS 1996

RESULT 13

AAU80079

ID AAU80079 standard; Protein; 123 AA.

XX

AC AAU80079;

XX

DT 30-JUL-2002 (first entry)

XX

DE Apoptin-associating protein 5 (AAP-5), partial sequence.

XX

KW Cancer; pACT; cytostatic; immunosuppressive; AAP-5;

KW Apoptin-associating protein 5; p53-independent apoptosis;

KW cell proliferation; cell death; autoimmune disease.

XX

OS Homo sapiens.

XX

PN EP1138768-A2.

XX

PD 04-OCT-2001.

XX

PF 27-MAR-2001; 2001EP-0201137.

XX

PR 27-MAR-2000; 2000EP-0201108.

XX

PA (LEAD-) LEADD BV.

XX

PI Noteborn MHM, Rohn JL, Danen-van Oorschot AAAM;

XX

DR WPI; 2001-657960/76.

XX

PT New apoptin-associating protein 5 for inducing p53-independent

PT apoptosis, or for treating cancer or autoimmune disease -

XX

PS Disclosure; Page 23; 44pp; English.

XX

CC The invention relates to an isolated or recombinant nucleic acid

CC which encodes an apoptin-associating protein 5 (AAP-5) capable of

CC providing apoptosis, or its functional fragment. The nucleic acid, a

CC vector comprising the nucleic acid, or a proteinaceous substance is

CC useful for the induction of p53-independent apoptosis. The pharmaceutical

CC composition comprising the nucleic acid, vector comprising the nucleic

CC acid or the proteinaceous substance is also useful for the induction of

CC p53-independent apoptosis, or for the treatment of a disease where

CC enhanced cell proliferation or decreased cell death is observed, such as

CC cancer or autoimmune disease. The present sequence represents the  
 CC partial amino acid sequence of apoptin-associating protein 5 (AAP-5).

XX SQ Sequence 123 AA;

Query Match 51.1%; Score 45; DB 22; Length 123;  
 Best Local Similarity 66.7%; Pred. No. 26;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYYSS 14  
 | | | | |  
 Db 8 YTSSQCGKYSS 19

RESULT 14  
 AAU80078  
 ID AAU80078 standard; Protein; 126 AA.

XX AC AAU80078;

XX 30-JUL-2002 (first entry)

XX DE Apoptin-associating protein 5 (AAP-5).

XX KW Cancer; pACT; cytostatic; immunosuppressive; AAP-5;  
 KW Apoptin-associating protein 5; p53-independent apoptosis;  
 KW cell proliferation; cell death; autoimmune disease.

XX OS Homo sapiens.

XX Key Location/Qualifiers  
 XX Region 1..3

XX FT /note= "Encoded by pACT vector"

XX EP1138768-A2.

XX PD 04-OCT-2001.

XX 27-MAR-2001; 2001EP-0201137.

XX 27-MAR-2000; 2000EP-0201108.

XX (LEAD-) LEADD BV.

XX Noteborn MHM, Rohn JL, Danen-van Oorschot AAAM;

XX WPI; 2001-657960/76.

XX N-PSDB; ABK30949.

XX New apoptin-associating protein 5 for inducing p53-independent  
 XX apoptosis, or for treating cancer or autoimmune disease

XX Claim 10; Fig 2; 44pp; English.

XX The invention relates to an isolated or recombinant nucleic acid  
 XX which encodes an apoptin-associating protein 5 (AAP-5) capable of  
 XX providing apoptosis, or its functional fragment. The nucleic acid, a  
 XX vector comprising the nucleic acid, or a proteinaceous substance is  
 XX useful for the induction of p53-independent apoptosis. The pharmaceuti-  
 XX cal composition comprising the nucleic acid, vector comprising the nucleic  
 XX acid or the proteinaceous substance is also useful for the induction of  
 XX p53-independent apoptosis, or for the treatment of a disease where  
 XX enhanced cell proliferation or decreased cell death is observed, such as  
 XX cancer or autoimmune disease. The present sequence represents the  
 XX amino acid sequence of apoptin-associating protein 5 (AAP-5).

XX SQ Sequence 126 AA;

Query Match 51.1%; Score 45; DB 22; Length 126;  
 Best Local Similarity 66.7%; Pred. No. 27;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYYSS 14

Db 11 YTSSQCGKYSS 22  
 | | | | |

RESULT 15

AA014806

ID AA014806 standard; Protein; 126 AA.

XX AC AA014806;

XX 08-JUL-2002 (first entry)

XX DE Human apoptin-associating protein 5, with pACT expression vector region.

XX KW Human; pACT expression vector; apoptin-associating protein 5; AAP-5;

XX AAP-6; p53-independent apoptosis; gene therapy; cancer;

XX KW autoimmune disease; apoptosis pathway cascade.

XX OS Homo sapiens.

XX Synthetic.

XX Key Location/Qualifiers

XX FT Peptide 1..3

XX FT /note= "This region is derived from the multiple cloning  
 FT site in the pACT expression vector"

XX FT 4..126

XX FT /note= "This region represents the human apoptin-  
 FT associating protein 5 (AAP-5)"

XX EP1138765-A1.

XX PD 04-OCT-2001.

XX 27-MAR-2000; 2000EP-0201108.

XX 27-MAR-2000; 2000EP-0201108.

XX (LEAD-) LEADD BV.

XX Noteborn MHM, Rohn JL, Danen-van Oorschot AAAM;

XX WPI; 2002-012523/02.

XX N-PSDB; AAL42576.

XX New proteins useful for inducing p53-independent apoptosis and for  
 XX treating cancer and autoimmune diseases comprises the isolated or  
 XX recombinant apoptin-associating proteinaceous substance

XX Disclosure; Fig 3; 28pp; English.

XX The invention comprises the amino acid and coding sequences of two  
 XX apoptin-associating proteins (AAP-5 and AAP-6). The AAP-5/AAP-6 DNA and  
 XX protein sequences are useful for inducing p53-independent apoptosis, and  
 XX the diagnosis/treatment (gene therapy) of cancer and autoimmune diseases.  
 XX The AAP-5/AAP-6 DNA and protein sequences are also useful for finding  
 XX additional apoptin-associating proteinaceous substances from the  
 XX apoptosis pathway cascade. The present amino acid sequence represents the  
 XX human AAP-5 protein (with a region derived from the multiple cloning  
 XX site of the pACT expression vector).

XX SQ Sequence 126 AA;

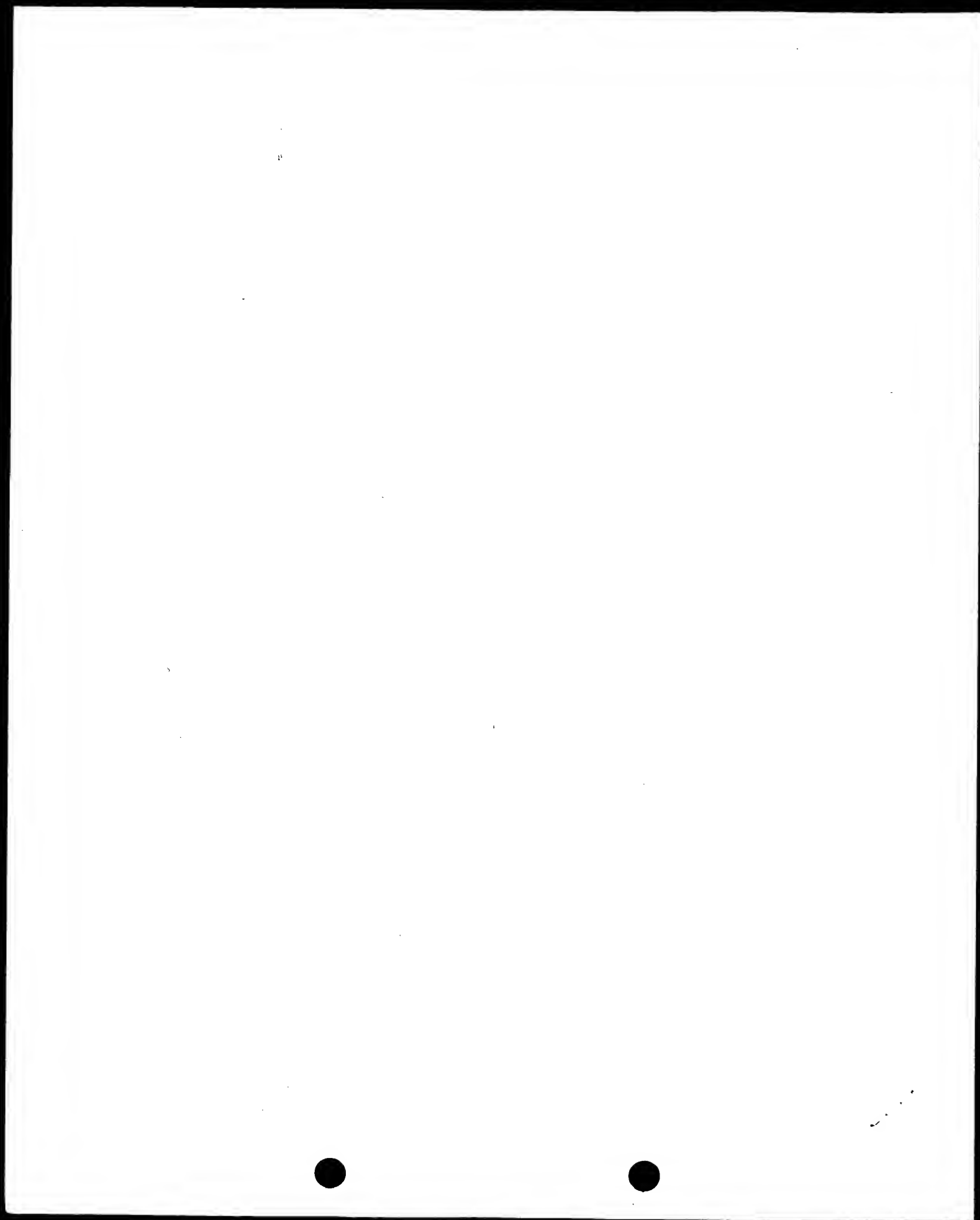
Query Match 51.1%; Score 45; DB 23; Length 126;  
 Best Local Similarity 66.7%; Pred. No. 27;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYYSS 14  
 | | | | |

Db 11 YTSSQCGKYSS 22

Search completed: December 2, 2002, 06:50:29  
 Job time: 4.11454 secs





GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:53:36 ; Search time 0.462555 Seconds  
(without alignments)  
516.399 Million cell updates/sec

Title: US-09-902-563-18  
Perfect score: 88  
Sequencing: 1 DRYPSGNGGLYSSG 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 102317

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2.6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2.6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2.6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2.6/ptodata/1/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2.6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2.6/ptodata/1/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2.6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2.6/ptodata/1/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2.6/ptodata/1/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2.6/ptodata/1/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2.6/ptodata/1/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2.6/ptodata/1/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2.6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2.6/ptodata/1/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Query Match | Score | Length | DB ID | Description         |
|------------|-------------|-------|--------|-------|---------------------|
| 1          | 45          | 51.1  | 126    | 10    | US-09-819-308-2     |
| 2          | 45          | 51.1  | 210    | 10    | US-09-819-308-10    |
| 3          | 40          | 45.5  | 316    | 12    | US-10-001-870-132   |
| 4          | 39          | 44.3  | 695    | 10    | US-09-764-898-191   |
| 5          | 39          | 44.3  | 837    | 12    | US-10-052-586-454   |
| 6          | 38.5        | 43.8  | 69     | 10    | US-09-925-300-1232  |
| 7          | 38.5        | 43.8  | 331    | 9     | US-09-344-882-6     |
| 8          | 38.5        | 43.8  | 406    | 9     | US-09-344-882-14    |
| 9          | 38          | 43.2  | 23     | 10    | US-09-785-632A-39   |
| 10         | 38          | 43.2  | 2485   | 10    | US-09-802-669-46    |
| 11         | 37.5        | 42.6  | 408    | 10    | US-09-864-761-37795 |
| 12         | 37          | 42.0  | 188    | 9     | US-10-125-452-21    |
| 13         | 37          | 42.0  | 303    | 12    | US-10-006-867-166   |
| 14         | 37          | 42.0  | 303    | 12    | US-10-052-586-562   |
| 15         | 37          | 42.0  | 321    | 10    | US-09-886-055-259   |
| 16         | 37          | 42.0  | 381    | 10    | US-09-815-242-5153  |
| 17         | 37          | 42.0  | 677    | 9     | US-10-125-452-34    |
| 18         | 37          | 42.0  | 778    | 9     | US-10-125-470-16    |
| 19         | 37          | 42.0  | 778    | 9     | US-10-125-452-16    |

20 37 42.0 918 9 US-10-020-733-4 Sequence 4, Appli  
21 37 42.0 918 9 US-10-125-470-9 Sequence 9, Appli  
22 37 42.0 918 9 US-10-125-452-9 Sequence 9, Appli  
23 37 42.0 926 9 US-09-983-531A-4 Sequence 4, Appli  
24 37 42.0 926 9 US-10-020-733-2 Sequence 2, Appli  
25 37 42.0 955 9 US-10-020-733-8 Sequence 8, Appli  
26 37 42.0 963 9 US-10-020-733-6 Sequence 6, Appli  
27 36.5 41.5 70 12 US-10-027-348-9 Sequence 9, Appli  
28 36.5 41.5 284 12 US-10-027-348-14 Sequence 14, Appli  
29 36.5 41.5 342 12 US-10-027-348-2 Sequence 2, Appli  
30 36.5 41.5 1066 9 US-09-423-126-3 Sequence 3, Appli  
31 36.5 41.5 1066 10 US-09-280-197-5 Sequence 5, Appli  
32 36 40.9 241 10 US-09-923-302-537 Sequence 537, App  
33 36 40.9 522 10 US-09-849-566-10 Sequence 10, Appl  
34 36 40.9 522 10 US-09-907-859-10 Sequence 10, Appl  
35 36 40.9 720 10 US-09-801-368-176 Sequence 176, App  
36 36 40.9 979 10 US-09-903-187A-5 Sequence 5, Appli  
37 35.5 40.3 485 9 US-09-860-846-10 Sequence 10, Appli  
38 35.5 40.3 485 10 US-09-861-289-10 Sequence 10, Appli  
39 35.5 40.3 3782 9 US-09-860-846-4 Sequence 4, Appli  
40 35.5 40.3 3782 10 US-09-861-289-4 Sequence 4, Appli  
41 35 39.8 10 10 US-09-819-308-25 Sequence 25, Appl  
42 35 39.8 52 10 US-09-864-761-35273 Sequence 35273, A  
43 35 39.8 57 10 US-09-764-877-1105 Sequence 1105, Ap  
44 35 39.8 77 10 US-09-864-761-34989 Sequence 34989, A  
45 35 39.8 220 10 US-09-462-846-3 Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-09-819-308-2  
; Sequence 2, Application US/09819308  
; Patent No. US20020019040A1  
; GENERAL INFORMATION:  
; APPLICANT: No. US20020019040Aleborn, Mathieu  
; APPLICANT: Danen-van Oorschot, Astrid  
; APPLICANT: Rohn, Jennifer  
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN  
; FILE REFERENCE: 2906-4820US  
; CURRENT APPLICATION NUMBER: US/09/819,308  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: vector pMT2SM-AAP-5  
US-09-819-308-2

Query Match 51.1%; Score 45; DB 10; Length 126;  
Best Local Similarity 66.7%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YPSGNGGLYSS 14  
| | | | |  
Db 11 YTSSCGKYYSS 22

RESULT 2  
US-09-819-308-10  
; Sequence 10, Application US/09819308  
; Patent No. US20020019040A1  
; GENERAL INFORMATION:  
; APPLICANT: No. US20020019040Aleborn, Mathieu  
; APPLICANT: Danen-van Oorschot, Astrid  
; APPLICANT: Rohn, Jennifer  
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN  
; FILE REFERENCE: 2906-4820US  
; CURRENT APPLICATION NUMBER: US/09/819,308  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 210  
; TYPE: PRT  
; ORGANISM: open reading frame of AAP-5  
US-09-819-308-10

Query Match 51.1%; Score 45; DB 10; Length 210;  
Best Local Similarity 66.7%; Pred. No. 5.7;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYSS 14  
Db 95 YTSQCCKYSS 106  
||| |||

RESULT 3  
US-10-001-870-132  
; Sequence 132, Application US/10001870  
; Patent No. US20020150924A1

GENERAL INFORMATION:

; APPLICANT: Salceda, Susana  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Herve  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
; FILE REFERENCE: DEX-0283  
; CURRENT APPLICATION NUMBER: US/10/001,870  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/252,189  
; PRIOR FILING DATE: 2000-11-21  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 132  
; LENGTH: 316  
; TYPE: PRT  
; ORGANISM: Homo sapien

US-10-001-870-132

Query Match 45.5%; Score 40; DB 12; Length 316;  
Best Local Similarity 75.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLYSSG 15  
Db 192 CGIYSPG 199  
||| |||

RESULT 4  
US-09-764-898-191  
; Sequence 191, Application US/09764898  
; Patent No. US20020050673A1  
GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PJ201  
; CURRENT APPLICATION NUMBER: US/09/764,898  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 311  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 191  
; LENGTH: 695  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-764-898-191

Query Match 44.3%; Score 39; DB 10; Length 695;  
Best Local Similarity 58.3%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 YPSGNCGLYSS 14

Db 321 YPGMCLPYST 332  
||| |||

RESULT 5

US-10-052-586-454  
; Sequence 454, Application US/10052586  
; Patent No. US20020127584A1  
GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C1  
; CURRENT FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063564  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063734  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/063870  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066120  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/066466  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/066772  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/069335  
; PRIOR FILING DATE: 1997-12-11  
; PRIOR APPLICATION NUMBER: 60/069425  
; PRIOR FILING DATE: 1997-12-12  
; PRIOR APPLICATION NUMBER: 60/069870  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/068017  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939,  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/083495  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083496  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083499  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083559  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/084366  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/086023  
; PRIOR FILING DATE: 1998-05-18  
; PRIOR APPLICATION NUMBER: 60/086392  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/086486  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087098  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087208  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088722  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088740  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088811  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088825  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088863  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089090  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089908

Query Match 44.3%; Score 39; DB 12; Length 837;  
Best Local Similarity 58.3%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 PSNCGLVYSSG 15

Db 522 PMANCSLRSCG 533

RESULT 6

US-09-925-300-1232

; Sequence 1232, Application US/09925300

```
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurttele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1232
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1232

Query Match 43.8%; Score 38.5; DB 10; Length 69;
Best Local Similarity 42.1%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 2; Indels 7; Gaps 1;

Qy 1 DRYPGSGN-----CGLYY 12
I: I I I I I I I I I I
Db 7 DKTPVGNPESGPGSCGLFY 25

RESULT 7
US-09-344-882-6
; Sequence 6, Application US/09344882
; Patent No. US20020162137A1
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurttele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinshan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allred, Carolyn C
; APPLICANT: Fatland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; FILE REFERENCE: 201573
; CURRENT APPLICATION NUMBER: US/09/344,882
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 6
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-09-344-882-6

Query Match 43.8%; Score 38.5; DB 9; Length 331;
Best Local Similarity 66.7%; Pred. No. 79;
Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 5 SGNCG-LYYSSG 15
I I I I I I I I I I
Db 102 SNNCGMLHYTSG 113

RESULT 8
US-09-344-882-14
; Sequence 14, Application US/09344882
; Patent No. US20020162137A1
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurttele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinshan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allred, Carolyn C
; APPLICANT: Fatland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; FILE REFERENCE: 201573
; CURRENT APPLICATION NUMBER: US/09/344,882
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 6
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-09-344-882-14

Query Match 43.2%; Score 38; DB 10; Length 23;
Best Local Similarity 54.5%; Pred. No. 7.6;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YPSGNGGLYY 13
I I I I I I I I I I
Db 1 YKGGCGKRFYS 11

RESULT 10
US-09-802-669-46
; Sequence 46, Application US/09802669
; Patent No. US2002000490A1
; GENERAL INFORMATION:
; APPLICANT: Nikolau, Basil J
; APPLICANT: Wurttele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
; APPLICANT: Ke, Jinshan
; APPLICANT: Johnson, Jerry L
; APPLICANT: Allred, Carolyn C
; APPLICANT: Fatland, Beth
; APPLICANT: Lutziger, Isabelle
; APPLICANT: Wen, Tsui-Jung
; TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
; FILE REFERENCE: 201573
; CURRENT APPLICATION NUMBER: US/09/344,882
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/090,717
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.2
; SEQ ID NO 14
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Arabidopsis Thaliana
US-09-344-882-14

Query Match 43.8%; Score 38.5; DB 9; Length 406;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 5 SGNCG-LYYSSG 15
I I I I I I I I I I
Db 177 SNNCGMLHYTSG 188

RESULT 9
US-09-785-632A-39
; Sequence 39, Application US/09785632A
; Patent No. US20020061512A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Kwon, Young Do
; APPLICANT: Kim, Hyun-Won
; APPLICANT: Ryu, Eun-Hyun
; APPLICANT: Hwang, Moon-Sun
; TITLE OF INVENTION: ZINC FINGER DOMAINS AND METHODS OF
; FILE REFERENCE: 12279-002001
; CURRENT APPLICATION NUMBER: US/09/785,632A
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-632A-39

Query Match 43.2%; Score 38; DB 10; Length 23;
Best Local Similarity 54.5%; Pred. No. 7.6;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YPSGNGGLYY 13
I I I I I I I I I I
Db 1 YKGGCGKRFYS 11

RESULT 10
US-09-802-669-46
; Sequence 46, Application US/09802669
; Patent No. US2002000490A1
; GENERAL INFORMATION:
```

APPLICANT: Dean, Nicholas M.  
APPLICANT: Marcussen, Eric G.  
APPLICANT: Wyatt, Jacqueline  
APPLICANT: Zhang, Hong  
TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling  
FILE REFERENCE: ISPH-545  
CURRENT APPLICATION NUMBER: US/09/802.669  
CURRENT FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: US/09/665.615  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/09/290.640  
PRIOR FILING DATE: 1999-04-12  
NUMBER OF SEQ ID NOS: 180  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 46  
LENGTH: 2485  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-802-669-46

Query Match 43.2%; Score 38; DB 10; Length 2485;  
Best Local Similarity 46.2%; Pred. No. 6.3e+02;  
Matches 6; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 YPSGNGGLYSSG 15  
||| | | | |  
Db 2019 YPKGKCTVOIKG 2031

## RESULT 11

US-09-864-761-37795  
; Sequence 37795, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Weisheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864.761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US/60/180.312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US/60/207.456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US/09/632.366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US/60/236.359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US/60/234.687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US/09/608.408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US/09/774.203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 37795  
; LENGTH: 408  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL049780.2  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.77  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.71  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1  
; OTHER INFORMATION: SWISSPROT HIT: P34618, EVALUATE 2.40e-01  
; OTHER INFORMATION: EST\_HUMAN HIT: BE390050.1, EVALUATE 1.00e-81  
US-09-864-761-37795

Query Match 42.6%; Score 37.5; DB 10; Length 408;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 9; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 1 DRYPGNGGLYSSG 15  
| | | | | | | | | |  
Db 80 DSGGQGN-GLFYSSG 93

## RESULT 12

US-10-125-452-21  
; Sequence 21, Application US/10125452  
; Patent No. US20020173640A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies  
; FILE REFERENCE: PT006P2  
; CURRENT APPLICATION NUMBER: US/10/125.452  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 09/955.504  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 09/712.907  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: PCT/US00/14308  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/178.717  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/142.930  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 60/136.388  
; PRIOR FILING DATE: 1999-05-27  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-125-452-21

Query Match 42.0%; Score 37; DB 9; Length 188;  
Best Local Similarity 60.0%; Pred. No. 77;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PSGNGGLYYS 13  
| | | | | | | | | |  
Db 155 PPGNGGFEHS 164

RESULT 13  
US-10-006-867-166

Sequence 166, Application US/10006867  
Patent No. US20020119130A1  
GENERAL INFORMATION:  
APPLICANT: Eaton, Dan L.  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3230R1C1  
CURRENT APPLICATION NUMBER: US/10/006,867  
CURRENT FILING DATE: 2001-12-06  
PRIOR APPLICATION NUMBER: 60/063435  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/064215  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088740  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088811  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088825  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088863  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090688  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091628  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/096012  
PRIOR FILING DATE: 1998-08-10  
PRIOR APPLICATION NUMBER: 60/096757  
PRIOR FILING DATE: 1998-08-17  
PRIOR APPLICATION NUMBER: 60/096949  
PRIOR FILING DATE: 1998-08-18  
PRIOR APPLICATION NUMBER: 60/096959  
PRIOR FILING DATE: 1998-08-18  
PRIOR APPLICATION NUMBER: 60/097954  
PRIOR FILING DATE: 1998-08-26  
PRIOR APPLICATION NUMBER: 60/097971  
PRIOR FILING DATE: 1998-08-26  
PRIOR APPLICATION NUMBER: 60/097979  
PRIOR FILING DATE: 1998-08-26  
PRIOR APPLICATION NUMBER: 60/098749  
PRIOR FILING DATE: 1998-09-01  
PRIOR APPLICATION NUMBER: 60/099741  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099763  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099792  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099812  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/099815  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/100627  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100662  
PRIOR FILING DATE: 1998-09-16  
PRIOR APPLICATION NUMBER: 60/100683  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/100684  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/100930  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/101279  
PRIOR FILING DATE: 1998-09-22  
PRIOR APPLICATION NUMBER: 60/101475  
PRIOR FILING DATE: 1998-09-23  
PRIOR APPLICATION NUMBER: 60/101738  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101743  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/101916  
PRIOR FILING DATE: 1998-09-24  
PRIOR APPLICATION NUMBER: 60/102570  
PRIOR FILING DATE: 1998-09-30  
PRIOR APPLICATION NUMBER: 60/103449  
PRIOR FILING DATE: 1998-10-06  
PRIOR APPLICATION NUMBER: 60/103678  
PRIOR FILING DATE: 1998-10-08  
PRIOR APPLICATION NUMBER: 60/103679  
PRIOR FILING DATE: 1998-10-08  
PRIOR APPLICATION NUMBER: 60/103711  
PRIOR FILING DATE: 1998-10-08  
PRIOR APPLICATION NUMBER: 60/105000  
PRIOR FILING DATE: 1998-10-20  
PRIOR APPLICATION NUMBER: 60/105002  
PRIOR FILING DATE: 1998-10-20  
PRIOR APPLICATION NUMBER: 60/105881  
PRIOR FILING DATE: 1998-10-27  
PRIOR APPLICATION NUMBER: 60/106030  
PRIOR FILING DATE: 1998-10-28  
PRIOR APPLICATION NUMBER: 60/106464  
PRIOR FILING DATE: 1998-10-30  
PRIOR APPLICATION NUMBER: 60/106856  
PRIOR FILING DATE: 1998-11-03  
PRIOR APPLICATION NUMBER: 60/108807  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: 60/112419  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/112422  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/112853  
PRIOR FILING DATE: 1998-12-16  
PRIOR APPLICATION NUMBER: 60/113011  
PRIOR FILING DATE: 1998-12-16  
PRIOR APPLICATION NUMBER: 60/112854

; PRIOR FILING DATE: 1998-12-16  
; PRIOR APPLICATION NUMBER: 60/113300  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/113408  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/113430  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113621  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/114223  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 60/115614  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/116527  
; PRIOR FILING DATE: 1999-01-20  
; PRIOR APPLICATION NUMBER: 60/116843  
; PRIOR FILING DATE: 1999-01-22  
; PRIOR APPLICATION NUMBER: 60/119285  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: 60/119287  
; PRIOR FILING DATE: 1999-02-09  
; PRIOR APPLICATION NUMBER: 60/119525  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/119549  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/120014  
; PRIOR FILING DATE: 1999-02-11  
; PRIOR APPLICATION NUMBER: 60/129122  
; PRIOR FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: 60/129674  
; PRIOR FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/131291  
; PRIOR FILING DATE: 1999-04-27  
; PRIOR APPLICATION NUMBER: 60/138387  
; PRIOR FILING DATE: 1999-06-09  
; PRIOR APPLICATION NUMBER: 60/144791  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: 60/169495  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: 60/175481  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 60/191007  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/199397  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/380139  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 09/311832  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/380137  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380138  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: 09/380142

Query Match 42.0%; Score 37; DB 12; Length 303;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 SGNCGIYSSG 15

DB 278 SGQCVIHSKG 288

## RESULT 14

US-10-052-586-562  
; Sequence 562, Application US/10052586  
; Patent No. US20020127584A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C1  
; CURRENT APPLICATION NUMBER: US/10/052,586  
; CURRENT FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063564  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063734  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/063870  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066120  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/066466  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/066772  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/069335  
; PRIOR FILING DATE: 1997-12-11  
; PRIOR APPLICATION NUMBER: 60/069425  
; PRIOR FILING DATE: 1997-12-12  
; PRIOR APPLICATION NUMBER: 60/069870  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/068017  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327

; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/083495  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083496  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083499  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083559  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/084366  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/086023  
; PRIOR FILING DATE: 1998-05-16  
; PRIOR APPLICATION NUMBER: 60/086392  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/086486  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087098  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087208  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088722  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088740  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088811  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088825  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088863  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089090  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089908

Query Match 42.0%; Score 37; DB 12; Length 303;

Best Local Similarity 54.5%; Pred. No. 1.2e+02; Mismatches 3; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 SGMGLVYSSG 15  
|||:||||  
Db 278 SGQCVIHYSG 288

#### RESULT 15

US-09-886-055-259  
; Sequence 259, Application US/09886055  
; Patent No. US2002013273A1  
; GENERAL INFORMATION:  
; APPLICANT: STRYER, LUBERT  
; APPLICANT: ZOZULYA, SERGEY  
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
; FILE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS  
; FILE REFERENCE: 078003-0277150  
; CURRENT APPLICATION NUMBER: US/09/886,055  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 60/213,812  
; PRIOR FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 522  
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID: NO 259
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-055-259
```

```
Query Match      42.0%; Score 37; DB 10; Length 321;
Best Local Similarity 85.7%; Pred. NO. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      6 GNCGLYY 12
        |||||
Db      43 GNCGLLY 49
```

```
Search completed: December 2, 2002, 07:00:07
Job time : 1.46256 secs
```

3  
4  
5

6  
7  
8

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:49:20 ; Search time 0.89207 seconds  
(without alignments)  
494.741 Million cell updates/sec

Title: US-09-902-563-18  
Perfect score: 88  
Sequence: 1 DRYPSGNGCLYSSG 15

Scoring Table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 88    | 100.0       | 15     | 4     | US-09-442-143A-18  |
| 2          | 88    | 100.0       | 432    | 4     | US-09-442-143A-4   |
| 3          | 88    | 100.0       | 439    | 4     | US-09-442-143A-2   |
| 4          | 46    | 52.3        | 666    | 4     | US-09-341-587-1    |
| 5          | 46    | 52.3        | 1785   | 4     | US-09-341-587-3    |
| 6          | 46    | 52.3        | 3460   | 4     | US-09-334-220-1    |
| 7          | 46    | 52.3        | 3461   | 4     | US-09-334-220-2    |
| 8          | 41.5  | 47.2        | 470    | 2     | US-08-933-821-6    |
| 9          | 41.5  | 47.2        | 470    | 3     | US-08-960-507-6    |
| 10         | 41.5  | 47.2        | 470    | 4     | US-09-136-828-6    |
| 11         | 41.5  | 47.2        | 470    | 4     | US-09-332-928A-6   |
| 12         | 41.5  | 47.2        | 470    | 4     | US-09-136-801-6    |
| 13         | 41.5  | 47.2        | 470    | 4     | US-09-332-929-6    |
| 14         | 41    | 46.6        | 39     | 1     | US-08-390-882A-1   |
| 15         | 41    | 46.6        | 39     | 1     | US-08-390-882A-2   |
| 16         | 40    | 45.5        | 1041   | 1     | US-08-494-714-2    |
| 17         | 40    | 45.5        | 1041   | 5     | PCT-US96-10782-2   |
| 18         | 39.5  | 44.9        | 99     | 2     | US-08-537-400-19   |
| 19         | 39.5  | 44.9        | 212    | 4     | US-09-106-568E-150 |
| 20         | 39.5  | 44.9        | 1391   | 4     | US-09-106-568E-8   |
| 21         | 39    | 44.3        | 38     | 1     | US-08-451-472-9    |
| 22         | 39    | 44.3        | 39     | 1     | US-08-390-882A-3   |
| 23         | 39    | 44.3        | 700    | 4     | US-09-408-647A-2   |
| 24         | 39    | 44.3        | 1077   | 4     | US-09-390-234-12   |
| 25         | 39    | 44.3        | 2254   | 2     | US-08-286-819A-28  |
| 26         | 39    | 44.3        | 2254   | 3     | US-08-980-357-28   |
| 27         | 38.5  | 43.8        | 406    | 4     | US-09-108-020-4    |

|    |      |      |      |   |                   |                   |
|----|------|------|------|---|-------------------|-------------------|
| 28 | 38.5 | 43.8 | 406  | 4 | US-09-108-020-42  | Sequence 42, Appl |
| 29 | 38   | 43.2 | 2465 | 2 | US-08-596-291-3   | Sequence 3, Appl  |
| 30 | 38   | 43.2 | 2465 | 3 | US-09-100-804-3   | Sequence 3, Appl  |
| 31 | 38   | 43.2 | 2466 | 3 | PCT-US94-09943-2  | Sequence 12, Appl |
| 32 | 38   | 43.2 | 2466 | 5 | US-09-290-640-46  | Sequence 2, Appl  |
| 33 | 38   | 43.2 | 2485 | 4 | US-09-290-640-46  | Sequence 46, Appl |
| 34 | 37   | 42.0 | 393  | 3 | US-09-127-124-1   | Sequence 1, Appl  |
| 35 | 37   | 42.0 | 393  | 4 | US-08-979-170-2   | Sequence 2, Appl  |
| 36 | 36.5 | 41.5 | 70   | 4 | US-09-165-827C-9  | Sequence 9, Appl  |
| 37 | 36.5 | 41.5 | 284  | 4 | US-09-165-827C-14 | Sequence 14, Appl |
| 38 | 36.5 | 41.5 | 342  | 4 | US-09-165-827C-2  | Sequence 2, Appl  |
| 39 | 36.5 | 41.5 | 549  | 3 | US-08-886-886-13  | Sequence 13, Appl |
| 40 | 36.5 | 41.5 | 551  | 3 | US-08-886-886-15  | Sequence 15, Appl |
| 41 | 36.5 | 41.5 | 1066 | 2 | US-08-633-770A-1  | Sequence 1, Appl  |
| 42 | 36   | 40.9 | 116  | 3 | US-09-184-658-48  | Sequence 48, Appl |
| 43 | 36   | 40.9 | 233  | 1 | US-07-890-422B-1  | Sequence 1, Appl  |
| 44 | 36   | 40.9 | 265  | 4 | US-09-199-637A-57 | Sequence 57, Appl |
| 45 | 36   | 40.9 | 289  | 3 | US-09-184-658-63  | Sequence 63, Appl |

ALIGNMENTS

RESULT 1  
US-09-442-143A-18  
; Sequence 18, Application US/09442143A  
; Patent No. 6403089  
; GENERAL INFORMATION:  
; APPLICANT: Levy, Gary  
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation  
; FILE REFERENCE: 9579-14  
; CURRENT APPLICATION NUMBER: US/09/442,143A  
; CURRENT FILING DATE: 1995-11-15  
; PRIOR APPLICATION NUMBER: US 60/046,537  
; PRIOR FILING DATE: 1997-05-17  
; PRIOR APPLICATION NUMBER: US 60/061,684  
; PRIOR FILING DATE: 1997-10-10  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 15  
; TYPE: PPT  
; ORGANISM: Homo sapiens  
US-09-442-143A-18

Query Match 100.0%; Score 88; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DRYPSGNGCLYSSG 15  
| | | | | | | | | | | | | | |  
Db 1 DRYPSGNGCLYSSG 15

RESULT 2  
US-09-442-143A-4  
; Sequence 4, Application US/09442143A  
; Patent No. 6403089  
; GENERAL INFORMATION:  
; APPLICANT: Levy, Gary  
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation  
; FILE REFERENCE: 9579-14  
; CURRENT APPLICATION NUMBER: US/09/442,143A  
; CURRENT FILING DATE: 1995-11-15  
; PRIOR APPLICATION NUMBER: US 60/046,537  
; PRIOR FILING DATE: 1997-05-17  
; PRIOR APPLICATION NUMBER: US 60/061,684  
; PRIOR FILING DATE: 1997-10-10  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4

```
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Murine fgl2
US-09-442-143A-4

Query Match      100.0%; Score 88; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
Db 357 DRYPSGNCGLYSSG 371

RESULT 3
US-09-442-143A-2
; Sequence 2, Application US/09442143A
; Patent No. 6403089
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-14
; CURRENT APPLICATION NUMBER: US/09/442,143A
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: US 60/046,537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061,684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Homo sapiens fgl2
US-09-442-143A-2

Query Match      100.0%; Score 88; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 9.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15
Db 364 DRYPSGNCGLYSSG 378

RESULT 4
US-09-341-587-1
; Sequence 1, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-341-587-1

Query Match      52.3%; Score 46; DB 4; Length 666;
Best Local Similarity 64.3%; Pred. No. 28;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 4 PSNGCG--LYSSG 15
Db 183 PSNGCGFLFYASG 196

RESULT 5
US-09-341-587-3
; Sequence 3, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1785
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-341-587-3

Query Match      52.3%; Score 46; DB 4; Length 1785;
Best Local Similarity 64.3%; Pred. No. 77;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 4 PSNGCG--LYSSG 15
Db 1134 PSSNCGGLFYASG 1147

RESULT 6
US-09-334-220-1
; Sequence 1, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/0F704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3460
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-334-220-1

Query Match      52.3%; Score 46; DB 4; Length 3460;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNCGLY--YSS 14
Db 1973 YPGNIGLYCPYSS 1986

RESULT 7
US-09-334-220-2
; Sequence 2, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
```

FILE REFERENCE: 2427/0F704  
CURRENT APPLICATION NUMBER: US/09/334, 220  
CURRENT FILING DATE: 1999-06-16  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 3461  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-334-220-2

Query Match 52.3%; Score 46; DB 4; Length 3461;  
Best Local Similarity 71.4%; Pred. No. 1.5e-02;  
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 YPSGNGLY--YSS 14  
||| ||| ||| |||  
Db 1974 YPGNGIYCPYSS 1987

ULT 8  
US-09-933-821-6  
Sequence 6, Application US/08933821  
Patent No. 5972338  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
TITLE OF INVENTION: Tie Ligands  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,821  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: P1130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-3216  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 470 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-933-821-6

Query Match 47.2%; Score 41.5; DB 2; Length 470;  
Best Local Similarity 60.0%; Pred. No. 94;  
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRPSGNGLYSSG 15  
| | | | | | | | | |  
Db 404 DSY-SGNCALYQRCG 417

RESULT 9  
US-08-960-507-6  
Sequence 6, Application US/08960507  
Patent No. 6057435  
GENERAL INFORMATION:

APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
TITLE OF INVENTION: Tie Ligands  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/960,507  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: P1130p1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-3216  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 470 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-960-507-6

Query Match 47.2%; Score 41.5; DB 3; Length 470;  
Best Local Similarity 60.0%; Pred. No. 94;  
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRPSGNGLYSSG 15  
| | | | | | | | | |  
Db 404 DSY-SGNCALYQRCG 417

RESULT 10  
US-09-136-828-6  
Sequence 6, Application US/09136828  
Patent No. 6350450  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
TITLE OF INVENTION: Tie Ligands  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,828  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: P1130RIA  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-3216  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 470 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-09-130-828-6

Query Match 47.2%; Score 41.5; DB 4; Length 470;  
Best Local Similarity 60.0%; Pred. No. 94;  
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15  
Db 404 DSY-SGNCALYQRG 417

## RESULT 11

US-09-332-928A-6  
Sequence 6, Application US/09332928A  
Patent No. 6368853

## GENERAL INFORMATION:

APPLICANT: Godowski, Paul J.  
Gurney, Austin L.  
TITLE OF INVENTION: Tie Ligands  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/332-928A  
FILING DATE: 14-Jun-1999  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/933,821  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: P1130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-3216  
TELEFAX: 650/952-9881

## INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 470 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-332-928A-6

Query Match 47.2%; Score 41.5; DB 4; Length 470;  
Best Local Similarity 60.0%; Pred. No. 94;  
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15  
Db 404 DSY-SGNCALYQRG 417

## RESULT 12

US-09-136-801-6  
Sequence 6, Application US/09136801

Patent No. 6413770  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J.  
Gurney, Austin L.  
Hillan, Kenneth  
APPLICANT: Botstein, David  
APPLICANT: Goddard, Audrey  
APPLICANT: Roy, Margaret  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Tumas, Daniel  
APPLICANT: Schwall, Ralph  
TITLE OF INVENTION: Tie Ligand Homologues  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,801  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: P1130P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-3216  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 470 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-09-136-801-6

Query Match 47.2%; Score 41.5; DB 4; Length 470;  
Best Local Similarity 60.0%; Pred. No. 94;  
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15  
Db 404 DSY-SGNCALYQRG 417

## RESULT 13

US-09-332-929-6  
Sequence 6, Application US/09332929  
Patent No. 6420542

## GENERAL INFORMATION:

APPLICANT: Godowski, Paul J.  
Gurney, Austin L.  
TITLE OF INVENTION: Tie Ligands  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatin (Genentech)

Query Match 47.2%; Score 41.5; DB 4; Length 470;  
Best Local Similarity 60.0%; Pred. No. 94;  
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 DRYPSGNCGLYSSG 15  
Db 404 DSY-SGNCALYQRG 417

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/332.929  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/933.821  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Ginger R.  
; REGISTRATION NUMBER: 33.055  
; REFERENCE/DOCKET NUMBER: P1130  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-3216  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 470 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; US-09-332-929-6  
Query Match 47.2%; Score 41.5; DB 4; Length 470;  
Best Local Similarity 60.0%; Pred. No. 94;  
Matches 9; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
Qy 1 DRYPSGNGGLYSSG 15  
Db 404 DSY-SGNCALYQGG 417  
RESULT 14  
US-08-390-882A-1  
; Sequence 1, Application US/08390882A  
; Patent No. 5688764  
; GENERAL INFORMATION:  
; APPLICANT: Kral, Robert M. Jr.; Krapcho, Karen; Johnson, Janice  
; TITLE OF INVENTION: Insecticidal Peptides from Spider Venom  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MADSON & METCALF  
; STREET: 950 FIRST INTERSTATE BUILDING  
; STREET: 170 SOUTH MAIN STREET  
; CITY: SALT LAKE CITY  
; STATE: UTAH  
; COUNTRY: USA  
; ZIP: 84101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/390.882A  
; FILING DATE: No. 5688764e assigned  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: L. CRAIG METCALF  
; REGISTRATION NUMBER: 31,398  
; REFERENCE/DOCKET NUMBER: 1094.2.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (801) 537-1700  
; TELEFAX: (801) 537-1799  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 39 amino acids  
; TYPE: amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL:  
; ANTI-SENSE: no  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM: Calisoga sp.  
; INDIVIDUAL ISOLATE: peptide B  
; ORGANELLE: Venom glands  
; US-08-390-882A-2  
Query Match 46.6%; Score 41; DB 1; Length 39;  
Best Local Similarity 40.0%; Pred. No. 9.1;  
Matches 8; Conservative 3; Mismatches 1; Indels 8; Gaps 1;  
Qy 2 RYP-----SGNCGLYS 13  
Db 5 RYPCSNKDCSCGNCGTFT 24

; ORGANISM: Calisoga sp.  
; INDIVIDUAL ISOLATE: peptide A  
; ORGANELLE: Venom glands  
; US-08-390-882A-1  
Query Match 46.6%; Score 41; DB 1; Length 39;  
Best Local Similarity 40.0%; Pred. No. 9.1;  
Matches 8; Conservative 3; Mismatches 1; Indels 8; Gaps 1;  
Qy 2 RYP-----SGNCGLYS 13  
Db 5 RYPCSNKDCSCGNCGTFT 24  
RESULT 15  
US-08-390-882A-2  
; Sequence 2, Application US/08390882A  
; Patent No. 5688764  
; GENERAL INFORMATION:  
; APPLICANT: Kral, Robert M. Jr.; Krapcho, Karen; Johnson, Janice  
; TITLE OF INVENTION: Insecticidal Peptides from Spider Venom  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MADSON & METCALF  
; STREET: 950 FIRST INTERSTATE BUILDING  
; STREET: 170 SOUTH MAIN STREET  
; CITY: SALT LAKE CITY  
; STATE: UTAH  
; COUNTRY: USA  
; ZIP: 84101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/390.882A  
; FILING DATE: No. 5688764e assigned  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: L. CRAIG METCALF  
; REGISTRATION NUMBER: 31,398  
; REFERENCE/DOCKET NUMBER: 1094.2.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (801) 537-1700  
; TELEFAX: (801) 537-1799  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 39 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL:  
; ANTI-SENSE: no  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM: Calisoga sp.  
; INDIVIDUAL ISOLATE: peptide B  
; ORGANELLE: Venom glands  
; US-08-390-882A-2  
Query Match 46.6%; Score 41; DB 1; Length 39;  
Best Local Similarity 40.0%; Pred. No. 9.1;  
Matches 8; Conservative 3; Mismatches 1; Indels 8; Gaps 1;  
Qy 2 RYP-----SGNCGLYS 13  
Db 5 RYPCSNKDCSCGNCGTFT 24  
Search completed: December 2, 2002, 06:54:36  
Job time : 2.89207 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: December 2, 2002, 07:00:11 ; Search time 55 seconds  
(without alignments)  
26.218 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRPSGNCGLYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 2520

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*\*

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID     | Description        |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1          | 31    | 35.2        | 14     | 2  | A58963 | alpha-conotoxin Cn |
| 2          | 31    | 35.2        | 15     | 1  | NRKAG  | alpha-conotoxin GI |
| 3          | 26    | 29.5        | 15     | 2  | P00017 | terminal protein - |
| 4          | 24    | 27.3        | 10     | 2  | C41946 | T-cell receptor ga |
| 5          | 23.5  | 26.7        | 12     | 2  | S25056 | Ig heavy chain - m |
| 6          | 23    | 26.1        | 10     | 2  | PT0243 | Ig heavy chain CRD |
| 7          | 23    | 26.1        | 13     | 1  | JTJG3  | tremorogen a-13 -  |
| 8          | 23    | 26.1        | 14     | 2  | PH0792 | T-cell receptor al |
| 9          | 23    | 26.1        | 15     | 2  | PH1366 | Ig heavy chain DJ  |
| 10         | 23    | 26.1        | 15     | 2  | PH0750 | T-cell receptor be |
| 11         | 23    | 26.1        | 15     | 2  | PU0110 | complement factor  |
| 12         | 22.5  | 25.6        | 11     | 2  | B41946 | T-cell receptor ga |
| 13         | 22    | 25.0        | 9      | 2  | S19329 | sperm-activating p |
| 14         | 22    | 25.0        | 11     | 2  | PT0302 | Ig heavy chain CRD |
| 15         | 22    | 25.0        | 13     | 1  | NRK2G  | alpha-conotoxin GI |
| 16         | 22    | 25.0        | 13     | 2  | T50173 | alpha-2 collagen - |
| 17         | 22    | 25.0        | 14     | 1  | NRKNIM | alpha-conotoxin MI |
| 18         | 22    | 25.0        | 14     | 2  | PH1305 | Ig heavy chain DJ  |
| 19         | 22    | 25.0        | 15     | 2  | C37765 | hypothetical prote |
| 20         | 21    | 23.9        | 9      | 2  | PT0288 | Ig heavy chain CRD |
| 21         | 21    | 23.9        | 9      | 2  | G41946 | T-cell receptor ga |
| 22         | 21    | 23.9        | 9      | 2  | PD0027 | pev-tachykinin - p |
| 23         | 21    | 23.9        | 10     | 2  | B38887 | T-cell receptor ga |
| 24         | 21    | 23.9        | 11     | 2  | PH1584 | Ig H chain V-D-J r |
| 25         | 21    | 23.9        | 12     | 2  | H41946 | T-cell receptor ga |
| 26         | 21    | 23.9        | 13     | 2  | A28953 | alpha-conotoxin SI |
| 27         | 21    | 23.9        | 13     | 2  | A61210 | antibiotic GE2270  |
| 28         | 21    | 23.9        | 15     | 2  | PT0085 | protein QA600027 - |
| 29         | 21    | 23.9        | 15     | 2  | S47387 | T-cell antigen rec |

30 21 23.9 15 2 S03955 acidic fibroblast  
31 21 23.9 15 2 S08301 epidermal growth f  
32 20 22.7 6 2 PT0652 T-cell receptor be  
33 20 22.7 11 2 S68637 acetylcholinestera  
34 20 22.7 13 2 PH1596 Ig H chain V-D-J r  
35 20 22.7 14 2 S58862 botulinum neurotox  
36 20 22.7 14 2 S58866 botulinum neurotox  
37 20 22.7 14 2 PT0252 Ig heavy chain CRD  
38 20 22.7 14 2 PH1597 Ig H chain V-D-J r  
39 19 21.6 7 2 A11483 aspartate transami  
40 19 21.6 10 2 A31571 hypertrehalosemic/  
41 19 21.6 12 2 S09082 proteasome chain 1  
42 19 21.6 12 2 PH0802 T-cell receptor al  
43 19 21.6 13 2 PH0788 T-cell receptor al  
44 19 21.6 14 2 PA0109 porin por 1B - Ara  
45 19 21.6 14 2 PA0045 porin por1 - Arabi

## ALIGNMENTS

### RESULT 1

A58963

alpha-conotoxin CnIA - cone shell (Conus consors)

N:Contains: alpha-conotoxin CnIB

C:Species: Conus consors

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999

C:Accession: A58963

R:Favreau, P.; Krimm, I.; Le Gall, F.; Bobenrieth, M.J.; Lamthanh, H.; Bouet, F.; Ser Biochemistry 38, 6317-6326, 1999

A>Title: Biochemical characterization and nuclear magnetic resonance structure of nov

A:Reference number: A58963; MUID:99255390; PMID:10320362

A:Accession: A58963

A>Status: Preliminary

A:Molecule type: protein

A:Residues: 1-14 <FAV>

C:Superfamily: alpha-conotoxin

C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neu

F:1-14/Product: alpha-conotoxin CnIA #status experimental <NATA>

F:3-14/Product: alpha-conotoxin CnIB #status experimental <NATB>

F:8-4-14/Disulfide bonds: #status experimental

F:14/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 35.2%; Score 31; DB 2; Length 14;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLIYYS 13

DB 8 CGKIYYS 13

### RESULT 2

NRKAG

alpha-conotoxin GIA [validated] - cone shell (Conus geographus)

N:Alternate names: alpha-CTX-GIA

N:Contains: alpha-conotoxin GI

C:Species: Conus geographus (geography cone)

C>Date: 24-Sep-1981 #sequence\_revision 24-Sep-1981 #text\_change 15-Sep-2000

C:Accession: A01782

R:Gray, W.R.; Luque, A.; Olivera, B.M.; Barrett, J.; Cruz, L.J.

J. Biol. Chem. 256, 4734-4740, 1981

A>Title: Peptide toxins from Conus geographus venom.

A:Reference number: A92320; MUID:81191854; PMID:7014556

A:Accession: A01782

A:Molecule type: protein

A:Residues: 1-15 <GRA>

R:Gray, W.R.; Rivier, J.E.; Galyean, R.; Cruz, L.J.; Olivera, B.M.

J. Biol. Chem. 256, 12247-12251, 1983

A>Title: Conotoxin MI. Disulfide bonding and conformational states.

A:Reference number: A92396; MUID:84032400; PMID:6630187

A:Contents: annotation; disulfide bonds

R:Guddat, L.W.; Shan, L.; Martin, J.L.; Edmundson, A.B.; Gray, W.R.

submitted to the Brookhaven Protein Data Bank, May 1996  
A:Reference number: A66253; PDB:1NO7  
A:Contents: annotation; X-ray crystallography, 1.2 angstroms, residues 1-13  
R:Gudat, L.W.; Martin, J.A.; Shan, L.; Edmundson, A.B.; Gray, W.R.  
Biochemistry 35, 11329-11335, 1996  
A:Title: Three-dimensional structure of the alpha-conotoxin GI at 1.2 angstroms resolution  
A:Reference number: A58592; MUID:96378624; PMID:8784187  
A:Contents: annotation; X-ray crystallography, 1.2 angstroms  
R:Pardi, A.; Galdes, A.; Florance, J.; Manicote, D.  
Biochemistry 28, 5494-5501, 1989  
A:Title: Solution structures of alpha-conotoxin GI determined by two-dimensional NMR spectroscopy  
A:Reference number: A30629; MUID:89375269; PMID:2775719  
A:Contents: annotation; conformation by (1)H-NMR  
C:Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynaptic antagonist  
C:Superfamily: alpha-conotoxin  
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; blocked carboxyl end  
F:1-15/Product: conotoxin GI #status experimental <GTA>  
F:1-13/Product: conotoxin GI #status experimental <GTA>  
F:2-7,3-13/Disulfide bonds: #link GIA #status predicted  
F:2-7,3-13/Disulfide bonds: #link GIC #status experimental  
F:13/Modified site: amidated carboxyl end (Cys) (amide in mature form from following glycosylation)  
F:15/Modified site: blocked carboxyl end (Lys) (probably amidated) #status experimental

Query Match 35.2%; Score 31; DB 1; Length 15;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYSSG 15  
DB 7 CGHYSG 14  
||| |||

RESULT 3  
P00017  
Terminal protein - phase M2 (fragment)  
C:Species: phase M2  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
C:Accession: P00017  
R:Matsumoto, K.; Takano, H.; Kim, C.I.; Hirokawa, H.  
Gene 84, 247-255, 1989  
A:Title: Primary structure of bacteriophage M2 DNA polymerase: conserved segments within the gene  
A:Reference number: JQ00161; MUID:90128268; PMID:2515115  
A:Accession: P00017  
A:Molecule type: DNA  
A:Residues: 1-15 <MAT>  
A:Cross-references: GB:M33144; NID:g215507; PIDN:AAA32367.1; PID:g215508  
C:Genetics:  
A:Gene: E  
C:Superfamily: phage P2A terminal protein

Query Match 29.5%; Score 26; DB 2; Length 15;  
Best Local Similarity 50.0%; Pred. No. 7.3e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 DRYPSGCG 10  
DB 1 DRYRGDNL 10  
||| |

RESULT 4  
C41946  
T-cell receptor gamma chain (1t.60) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
C:Accession: C41946  
R:Whetzel, M.; Mosley, R.L.; Whetzel, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.  
Mol. Cell. Biol. 11, 5902-5909, 1991  
A:Title: Rearrangement and functional-site sequence analyses of T-cell receptor gamma gene  
A:Reference number: A41946; MUID:92049316; PMID:1658619  
A:Accession: C41946  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-10 <WHE>

C:Keywords: T-cell receptor

Query Match 27.3%; Score 24; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1e+03;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYSSG 15  
DB 2 CAVSSSG 9  
||| |||

RESULT 5  
S25056  
Ig heavy chain - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 25-Feb-1994 #sequence\_revision 01-Dec-1995 #text\_change 20-Jun-2000  
C:Accession: S25056  
R:Jacob, J.; Kelsoe, G.  
submitted to the EMBL Data Library, July 1992  
A:Description: In situ studies on the primary immune response to (4-hydroxy-3-nitrophenyl)antigen  
A:Reference number: S25024  
A:Accession: S25056  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-12 <JAC>  
A:Cross-references: EMBL:X67386; NID:g50927; PIDN:CAA47798.1; PID:g1333920  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.7%; Score 23.5; DB 2; Length 12;  
Best Local Similarity 46.2%; Pred. No. 1.5e+03;  
Matches 6; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 2 RYPSGCGLYSS 14  
DB 2 RYP-----YYGS 9  
||| |||

RESULT 6  
PT0243  
Ig heavy chain CRD3 region (clone 2-103A) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0243  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an  
A:Reference number: PT0222; MUID:91108337; PMID:1899102  
A:Accession: PT0243  
A:Molecule type: DNA  
A:Residues: 1-10 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.1%; Score 23; DB 2; Length 10;  
Best Local Similarity 71.4%; Pred. No. 1.5e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 GLYSSG 15  
DB 2 GYDSSG 8  
||| |||

RESULT 7  
JTJG3  
Tremorgen a-13 - jelly fungus (Tremella mesenterica)  
C:Species: Tremella mesenterica  
C:Date: 03-Aug-1984 #sequence\_revision 03-Aug-1984 #text\_change 07-Nov-1997  
C:Accession: A01641  
R:Sakagami, Y.; Yoshida, M.; Isogai, A.; Suzuki, A.  
Science 212, 1525-1527, 1981  
A:Title: Peptide sex hormones inducing conjugation tube formation in compatible mating  
A:Reference number: A94256

A:Accession: A01641  
 A:Molecule type: protein  
 A:Residues: 1-13 <SAK>  
 C:Comment: Tremorogen a-13 is produced by the a mating-type cells and induces formation  
 C:Superfamily: tremorogen a-13  
 C:Keywords: extracellular protein; hormone; lipoprotein; pheromone; pranylated cysteine  
 F;13/Binding site: farnesyl (Cys) (covalent) #status experimental  
 F;13/Modified site: methyl ester carboxyl end (Cys) #status absent

Query Match 26.1%; Score 23; DB 1; Length 13;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PSGNC 8  
 Db 9 PSGVC 13

RESULT 8  
 PH0792  
 T-cell receptor alpha chain (H2 V-alpha-2.TA19) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PH0792  
 R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
 J. Exp. Med. 174, 1371-1383, 1991  
 A:Title: T cell receptor genes in a series of class I major histocompatibility complex-1  
 allelic exclusion and antigen-specific repertoire.  
 A:Reference number: PH0746; MUID:92078846; PMID:1836010  
 A:Accession: PH0792  
 A:Molecule type: mRNA  
 A:Residues: 1-14 <CAS>  
 A:Cross-references: EMBL:X60897  
 A:Experimental source: T lymphocyte  
 C:Keywords: T-cell receptor

Query Match 26.1%; Score 23; DB 2; Length 14;  
 Best Local Similarity 45.5%; Pred. No. 2.1e+03;  
 Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 SGNCGLYSSG 15  
 Db 4 SANSCTYORFG 14

RESULT 9  
 PH1366  
 Ig heavy chain DJ region (clone C111-106) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
 C:Accession: PH1366  
 R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.  
 J. Exp. Med. 176, 1577-1581, 1992  
 A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph  
 A:Reference number: PH1302; MUID:93094761; PMID:1460419  
 A:Accession: PH1366  
 A:Molecule type: DNA  
 A:Residues: 1-15 <WAS>  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 26.1%; Score 23; DB 2; Length 15;  
 Best Local Similarity 45.5%; Pred. No. 2.2e+03;  
 Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 SGNCGLYSSG 15  
 Db 4 SGYGDYWGQ 14

RESULT 10  
 PH0750  
 T-cell receptor beta chain (C11) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PH0750  
 R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
 J. Exp. Med. 174, 1371-1383, 1991  
 A:Title: T cell receptor genes in a series of class I major histocompatibility complex  
 allelic exclusion and antigen-specific repertoire.  
 A:Reference number: PH0746; MUID:92078846; PMID:1836010  
 A:Accession: PH0750  
 A:Molecule type: mRNA  
 A:Residues: 1-15 <CAS>  
 A:Cross-references: EMBL:X06841  
 A:Experimental source: T lymphocyte  
 C:Keywords: T-cell receptor

Query Match 26.1%; Score 23; DB 2; Length 15;  
 Best Local Similarity 44.4%; Pred. No. 2.2e+03;  
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 PSGNCGLY 12  
 Db 5 PTANTGQLY 13

RESULT 11  
 PLO110  
 Complement factor B1-Bb and B2-Bb - guinea pig (fragment)  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 15-Nov-1996  
 C:Accession: PLO110  
 R:Matsumura, M.; Okada, H.  
 Mol. Immunol. 26, 669-676, 1989  
 A:Title: Two forms of guinea pig factor B of the alternative complement pathway with  
 A:Reference number: A93136; MUID:89384686; PMID:2779589  
 A:Accession: PLO110  
 A:Molecule type: protein  
 A:Residues: 1-15 <MAT>  
 C:Keywords: complement alternate pathway; glycoprotein

Query Match 26.1%; Score 23; DB 2; Length 15;  
 Best Local Similarity 37.5%; Pred. No. 2.2e+03;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSGNCGLY 11  
 Db 6 PAGSMNTY 13

RESULT 12  
 B41946  
 T-cell receptor gamma chain (lt.57) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
 C:Accession: B41946  
 R:Whetzel, M.; Mosley, R.L.; Whetzel, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.  
 Mol. Cell. Biol. 11, 5902-5909, 1991  
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma  
 A:Reference number: A41946; MUID:92049316; PMID:1658619  
 A:Accession: B41946  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-11 <WHE>  
 C:Keywords: T-cell receptor

Query Match 25.6%; Score 22.5; DB 2; Length 11;  
 Best Local Similarity 55.6%; Pred. No. 2e+03;  
 Matches 5; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 8 CGLY-YSSG 15  
 Db 2 CAVWRYSSG 10

RESULT 13

## S19329

sperm-activating peptide SAP - sea urchin (Stomopneustes variolus)  
 C:Species: Stomopneustes variolus  
 C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 18-Aug-2000  
 C:Accession: S19329  
 R:Koshino, K.; Takao, T.; Shimonishi, Y.; Suzuki, N.  
 FEBS Lett. 294, 179-182, 1991  
 A:Title: Determination of the amino acid sequence of an intramolecular disulfide linkage  
 A:Reference number: S19329; MUID:92097763; PMID:1756858  
 A:Accession: S19329  
 A:Molecule type: protein  
 A:Residues: 1-9 <YOS>  
 C:Superfamily: unassigned animal peptides  
 F:3-8/Disulfide bonds: #status predicted

Query Match 25.0%; Score 22; DB 2; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSGNC 8  
 I I I I  
 Db 4 PEGKC 8

RESULT 14  
PT0302

Ig heavy chain CRD3 region (clone 5-112) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0302  
 R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
 A:Reference number: PT0222; MUID:91108337; PMID:1899102  
 A:Accession: PT0302  
 A:Molecule type: DNA  
 A:Residues: 1-11 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 25.0%; Score 22; DB 2; Length 11;  
 Best Local Similarity 71.4%; Pred. No. 2.4e+03;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 GLYSSG 15  
 I I I I I I  
 Db 3 GPRYSSG 9

RESULT 15  
NPKN2G

alpha-conotoxin GII - cone shell (Conus geographus)  
 C:Species: Conus geographus (geography cone)  
 C:Date: 24-Sep-1981 #sequence\_revision 24-Sep-1981 #text\_change 23-May-1997  
 C:Accession: A01783  
 R:Gray, W.R.; Luque, A.; Olivera, B.M.; Barrett, J.; Cruz, L.J.  
 J. Biol. Chem. 256, 4734-4740, 1981  
 A:Title: Peptide toxins from Conus geographus venom.  
 A:Reference number: A92320; MUID:81191854; PMID:7014556  
 A:Accession: A01783  
 A:Molecule type: protein  
 A:Residues: 1-13 <GRA>  
 R:Gray, W.R.; Rivier, J.E.; Galyean, R.; Cruz, L.J.; Olivera, B.M.  
 J. Biol. Chem. 258, 12247-12251, 1983  
 A:Title: Conotoxin MI. Disulfide bonding and conformational states.  
 A:Reference number: A92396; MUID:84032400; PMID:6630187  
 C:Contents: annotation: disulfide bonds  
 C:Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynapt  
 C:Superfamily: alpha-conotoxin  
 C:Keywords: acetylcholine receptor inhibitor; blocked carboxyl end: postsynaptic neuroto  
 F:2-7,3-13/Disulfide bonds: #status predicted  
 F:13/Modified site: blocked carboxyl end (Cys) (probably amidated) #status experimental

Query Match 25.0%; Score 22; DB 1; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLYYS 13  
 I I I I I  
 Db 7 CGKHFS 12

Search completed: December 2, 2002, 07:03:59  
 Job time : 56 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:54:40 ; Search time 11 seconds  
(without alignments)  
56.559 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRYPSGNCLYYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 748

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 31    | 35.2        | 14     | 1     | CXAL_CONCN  |
| 2          | 31    | 35.2        | 15     | 1     | CXAL_CONGE  |
| 3          | 26    | 29.5        | 15     | 1     | TERM_BPW2   |
| 4          | 23    | 26.1        | 13     | 1     | TAI3_TREME  |
| 5          | 23    | 26.1        | 14     | 1     | MARI_ALTPSP |
| 6          | 23    | 26.1        | 15     | 1     | UC14_MAIZE  |
| 7          | 22    | 25.0        | 9      | 1     | SAP_STOVA   |
| 8          | 22    | 25.0        | 13     | 1     | CXAL_CONGE  |
| 9          | 22    | 25.0        | 14     | 1     | CXAL_CONMA  |
| 10         | 21    | 23.9        | 9      | 1     | TRP4_LEUMA  |
| 11         | 21    | 23.9        | 15     | 1     | CLOA_FAT    |
| 12         | 21    | 23.9        | 15     | 1     | FGF1_CANFA  |
| 13         | 20    | 22.7        | 10     | 1     | PKCK_FASHE  |
| 14         | 19    | 21.6        | 9      | 1     | TKL1_LOCOMI |
| 15         | 19    | 21.6        | 10     | 1     | HTF_HELZE   |
| 16         | 19    | 21.6        | 14     | 1     | SCK3_LEIOU  |
| 17         | 19    | 21.6        | 15     | 1     | TA1_TREBR   |
| 18         | 19    | 21.6        | 15     | 1     | UC06_MAIZE  |
| 19         | 18    | 20.5        | 10     | 1     | AMPN_HELAM  |
| 20         | 18    | 20.5        | 13     | 1     | URAL_HUMAN  |
| 21         | 18    | 20.5        | 14     | 1     | CAT2_FASHE  |
| 22         | 18    | 20.5        | 14     | 1     | HY14_PIG    |
| 23         | 18    | 20.5        | 14     | 1     | LPER_BACLI  |
| 24         | 18    | 20.5        | 15     | 1     | KLOM_LUMTE  |
| 25         | 18    | 20.5        | 15     | 1     | UC19_MAIZE  |
| 26         | 17    | 19.3        | 8      | 1     | ALL1_CYPDO  |
| 27         | 17    | 19.3        | 8      | 1     | LCK8_LEUMA  |
| 28         | 17    | 19.3        | 9      | 1     | OXYT_RABIT  |
| 29         | 17    | 19.3        | 9      | 1     | OXYT_RABIT  |
| 30         | 17    | 19.3        | 12     | 1     | CXAL_CONMR  |
| 31         | 17    | 19.3        | 14     | 1     | SODN_STRGR  |
| 32         | 17    | 19.3        | 15     | 1     | OBPA_MAMBR  |
| 33         | 16    | 18.2        | 8      | 1     | ALL6_CYPDO  |

|    |    |      |    |   |             |                     |
|----|----|------|----|---|-------------|---------------------|
| 34 | 16 | 18.2 | 9  | 1 | COW_CONVE   | P83047 conus ventr  |
| 35 | 16 | 18.2 | 10 | 1 | COXO_THUOB  | P80982 thunnus obe  |
| 36 | 16 | 18.2 | 10 | 1 | CU30_LOCOMI | P11735 locusta mig  |
| 37 | 16 | 18.2 | 11 | 1 | COR2_PERAM  | P11496 periplaneta  |
| 38 | 16 | 18.2 | 11 | 1 | UN05_CLOPA  | P81350 clostridium  |
| 39 | 16 | 18.2 | 12 | 1 | FARI_CALVO  | P41869 calliphora   |
| 40 | 16 | 18.2 | 12 | 1 | FRE1_LITIN  | P82021 litorea inf  |
| 41 | 16 | 18.2 | 12 | 1 | LICH_BACLI  | P82907 bacillus li  |
| 42 | 16 | 18.2 | 12 | 1 | REF_CONSP   | P58805 conus spuri  |
| 43 | 16 | 18.2 | 13 | 1 | LPAA_PORGI  | P81411 porphyromon  |
| 44 | 16 | 18.2 | 13 | 1 | WPI1_MICOC  | P81532 micropililis |
| 45 | 16 | 18.2 | 14 | 1 | ADYT_ALYOB  | P08944 alytes obst  |

## ALIGNMENTS

|  |   |           |      |        |  |  |
|--|---|-----------|------|--------|--|--|
| RESULT 1   |   |           |      |        |  |  |
| CXAL_CONCN   |   |           |      |        |  |  |
| ID   | CXAL_CONCN  | STANDARD; | PRT; | 14 AA. |  |  |
| AC   | P56973;   |           |      |        |  |  |
| DT   | 16-OCT-2001 (Rel. 40, Created)  |           |      |        |  |  |
| DT   | 16-OCT-2001 (Rel. 40, Last sequence update)                           |           |      |        |  |  |
| DT   | 15-JUN-2002 (Rel. 41, Last annotation update)                         |           |      |        |  |  |
| DE   | Alpha-conotoxin CN1A [Contains: Alpha-conotoxin CN1B].                |           |      |        |  |  |
| OS   | Conus consors (Singed cone).  |           |      |        |  |  |
| OC   | Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;            |           |      |        |  |  |
| OC   | Neogastropoda; Conoidea; Conidae; Conus.                              |           |      |        |  |  |
| OX   | NCBI_TaxID=101297;  |           |      |        |  |  |
| RN   | [1]   |           |      |        |  |  |
| RP   | SEQUENCE, SYNTHESIS, CHARACTERIZATION, AND STRUCTURE BY NMR.          |           |      |        |  |  |
| RC   | TISSUE=Venom;   |           |      |        |  |  |
| RX   | MEDLINE=95255390; PubMed=10320362;                                    |           |      |        |  |  |
| RA   | Favreau P., Krimm I., le Gall F., Bobenrieth M.J., Lamthanh H.,       |           |      |        |  |  |
| RA   | Bouet F., Servent D., Molgo J., Menez A., Letourneux Y.,              |           |      |        |  |  |
| RA   | Lancelin J.-M.;   |           |      |        |  |  |
| RT   | "Biochemical characterization and nuclear magnetic resonance          |           |      |        |  |  |
| RT   | structure of novel alpha-conotoxins isolated from the venom of Conus  |           |      |        |  |  |
| RT   | consors."   |           |      |        |  |  |
| RL   | Biochemistry 38:6317-6326(1999).                                      |           |      |        |  |  |
| CC   | !- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY     |           |      |        |  |  |
| CC   | BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS        |           |      |        |  |  |
| CC   | INHIBIT THEM. THIS PEPTIDE SEEMS TO BE A POTENT AND SELECTIVE         |           |      |        |  |  |
| CC   | BLOCKER OF MUSCULAR SUBTYPE OF NACHR.                                 |           |      |        |  |  |
| CC   | !- SUBCELLULAR LOCATION: Secreted.                                    |           |      |        |  |  |
| CC   | !- TISSUE SPECIFICITY: Expressed by the venom duct.                   |           |      |        |  |  |
| CC   | !- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE |           |      |        |  |  |
| CC   | FAMILY.   |           |      |        |  |  |
| PDB:   | 1B45; 09-JUL-99.  |           |      |        |  |  |
| DR   | Postsynaptic neurotoxin; Neurotoxin; Toxin;                           |           |      |        |  |  |
| KW   | Acetylcholine receptor inhibitor; Amidation; 3D-structure.            |           |      |        |  |  |
| FT   | PEPTIDE 1 14  |           |      |        |  |  |
| FT   | PEPTIDE 3 14  |           |      |        |  |  |
| FT   | DISULFID 3 8  |           |      |        |  |  |
| FT   | DISULFID 4 14   |           |      |        |  |  |
| FT   | MOD_RES 14 14   |           |      |        |  |  |
| SQ   | SEQUENCE 14 AA; 1548 MW; DEEE1969BF5E5BD CRC64;                       |           |      |        |  |  |
| AMIDATION.   |   |           |      |        |  |  |
| Query Match 35.2%; Score 31; DB 1; Length 14;              |   |           |      |        |  |  |
| Best Local Similarity 83.3%; Pred. No. 44;                 |   |           |      |        |  |  |
| Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0; |   |           |      |        |  |  |
| QY   | 8   | CGLYYS 13 |      |        |  |  |
|  |   |           |      |        |  |  |
| Db   | 8   | CGKYYS 13 |      |        |  |  |
| RESULT 2   |   |           |      |        |  |  |
| CXAL_CONGE   |   |           |      |        |  |  |
| ID   | CXAL_CONGE  | STANDARD; | PRT; | 15 AA. |  |  |
| AC   | P01519;   |           |      |        |  |  |
| DT   | 21-JUL-1986 (Rel. 01, Created)  |           |      |        |  |  |
| DT   | 21-JUL-1986 (Rel. 01, Last sequence update)                           |           |      |        |  |  |

15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Alpha-conotoxin GIA [Contains: Alpha-conotoxin GI (GI)].  
 OS Conus geographus (Geography cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conidae; Conidae; Conus.  
 OX NCBI\_TaxID=6491;  
 RN (1)  
 RN SEQUENCE.  
 RP MEDLINE=81191854; PubMed=7014556;  
 RX Gray W.R., Luque F.A., Olivera B.M., Barrett J., Cruz L.J.;  
 RA "Peptide toxins from Conus geographus venom.";  
 RT J. Biol. Chem. 256:4734-4740(1981).  
 RL (2)  
 RN DISULFIDE BONDS OF GI, AND SYNTHESIS OF GI.  
 RP PubMed=7152021;  
 RX Nishiuchi Y., Sakakibara S.;  
 RA "Primary and secondary structure of conotoxin GI, a neurotoxic  
 RT tridecapeptide from a marine snail.";  
 RT FEBS Lett. 148:260-262(1982).  
 RL (3)  
 RN DISULFIDE BONDS OF GI, AND SYNTHESIS OF GI.  
 RP PubMed=6466616;  
 RX Gray W.R., Luque F.A., Galyean R., Atherton E., Sheppard R.C.,  
 RA Stone B.L., Reyes A., Alford J., McIntosh M., Olivera B.M.,  
 RA Cruz L.J., Rivier J.;  
 RT "Conotoxin GI: disulfide bridges, synthesis, and preparation of  
 RT iodinated derivatives.";  
 RL Biochemistry 23:2796-2802(1984).  
 RL (4)  
 RN COMPARISON WITH ALPHA-CONOTOXIN SI AND ALPHA-CONOTOXIN MI.  
 RP PubMed=7947815;  
 RX Hann R.M., Pagan O.R., Eterovic V.A.;  
 RA "The alpha-conotoxins GI and MI distinguish between the nicotinic  
 RT acetylcholine receptor agonist sites while SI does not.";  
 RL Biochemistry 33:14058-14063(1994).  
 RL (5)  
 RN PHARMACOLOGICAL CHARACTERIZATION ON MOUSE MUSCLE-DERIVED BC3H-1 CELLS  
 RP AND TORPEDO ELECTRIC ORGAN.  
 RX PubMed=7623764;  
 RA Groebe D.R., Dumm J.M., Levitan E.S., Abramson S.N.;  
 RT "Alpha-Conotoxins selectively inhibit one of the two acetylcholine  
 RT binding sites of nicotinic receptors.";  
 RL Mol. Pharmacol. 48:105-111(1995).  
 RL (6)  
 RN MUTAGENESIS OF ARG-9.  
 RP PubMed=9174364;  
 RX Groebe D.R., Gray W.R., Abramson S.N.;  
 RA "Determinants involved in the affinity of alpha-conotoxins GI and SI  
 RT for the muscle subtype of nicotinic acetylcholine receptors.";  
 RL Biochemistry 36:6469-6474(1997).  
 RL (7)  
 RN X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS) OF GI.  
 RP MEDLINE=96378624; PubMed=8784187;  
 RX Guddat L.W., Martin J.A., Shan L., Edmondson A.B., Gray W.R.;  
 RA "Three-dimensional structure of the alpha-conotoxin GI at 1.2-A  
 RT resolution.";  
 RL Biochemistry 35:11329-11335(1996).  
 RL (8)  
 RN STRUCTURE BY NMR OF GI.  
 RP PubMed=2765514;  
 RX Kobayashi Y., Ohkubo T., Kyogoku Y., Nishiuchi Y., Sakakibara S.,  
 RA Braun W., Go N.;  
 RT "Solution conformation of conotoxin GI determined by 1H nuclear  
 RT magnetic resonance spectroscopy and distance geometry calculations.";  
 RL Biochemistry 28:4853-4860(1989).  
 RL (9)  
 RN STRUCTURE BY NMR OF GI.  
 RP MEDLINE=89375269; PubMed=2775719;  
 RX Pardi A., Galdes A., Florance J., Manicote D.;  
 RA "Solution structures of alpha-conotoxin GI determined by two-  
 RT dimensional NMR spectroscopy.";  
 RL Biochemistry 28:5494-5501(1989).  
 RL (10)

RP STRUCTURE BY NMR OF GI.  
 RX PubMed=9660176;  
 RA Maslennikov I.V., Sobol A.G., Gladky K.V., Lugovskoy A.A.,  
 RA Ostrovsky A.G., Tsetlin V.I., Ivanov V.T., Arseniev A.S.;  
 RT "Two distinct structures of alpha-conotoxin GI in aqueous solution.";  
 RL Eur. J. Biochem. 254:238-247(1998).  
 RN (11)  
 RN STRUCTURE BY NMR OF GI.  
 RP MEDLINE=98239743; PubMed=9571060;  
 RX Gehrmann J., Alewood P.F., Craik D.J.;  
 RA "Structure determination of the three disulfide bond isomers of  
 RT alpha-conotoxin GI: a model for the role of disulfide bonds in  
 RT structural stability.";  
 RL J. Mol. Biol. 278:401-415(1998).  
 RN (12)  
 RN STRUCTURE BY NMR OF AN ANTITOXIC ANALOG OF GI.  
 RX MEDLINE=99438341; PubMed=10508392;  
 RA Mok K.H., Han K.H.;  
 RT "NMR solution conformation of an antitoxic analogue of alpha-conotoxin  
 RT GI: identification of a common nicotinic acetylcholine receptor  
 RT alpha(1)-subunit binding surface for small ligands and alpha-  
 RT conotoxins.";  
 RL Biochemistry 38:11895-11904(1999).  
 CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they  
 CC bind to the nicotinic acetylcholine receptors (nAChR) and thus  
 CC inhibit them. The higher affinity site for alpha-conotoxin GI is  
 CC the alpha/delta site on mouse muscle-derived BC3H-1 receptor, and  
 CC the other site (alpha/gamma site) on nicotinic receptors from  
 CC Torpedo californica electric organ.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
 CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE  
 CC FAMILY.  
 DR PIR: A01782; WTKNAG.  
 DR PDB: INOT; 07-DEC-96.  
 DR PDB: 1XGA; 16-FEB-99.  
 DR PDB: 1XGB; 16-FEB-99.  
 DR PDB: 1XGC; 23-MAR-99.  
 DR PDB: 1Q53; 06-OCT-99.  
 DR PDB: 1Q53; 06-OCT-99.  
 KW Postsynaptic neurotoxin; Neurotoxin; Toxin;  
 KW Acetylcholine receptor inhibitor; Amidation; 3D-structure.  
 FT PEPTIDE 1 15 ALPHA-CONOTOXIN GIA..  
 FT PEPTIDE 1 13 ALPHA-CONOTOXIN GI..  
 FT DISULFID 2 7  
 FT DISULFID 3 13  
 FT MOD\_RES 13 13  
 FT MOD\_RES 15 15  
 FT MUTAGEN 9 9  
 FT AMIDATION (G-14 PROVIDE AMIDE GROUP) (IN  
 FT GI).  
 FT AMIDATION (IN GIA).  
 FT R->A: REDUCTION IN AFFINITY FOR BOTH  
 FT ALPHA/DELTA AND ALPHA/GAMMA SITES ON  
 FT BC3H-1 RECEPTORS AND LOSS OF AFFINITY FOR  
 FT BOTH ALPHA/DELTA AND ALPHA/GAMMA SITES ON  
 FT TORPEDO RECEPTORS (IN GI).  
 FT 2AE73EE90F8C2E19 CRC64;  
 SQ SEQUENCE 15 AA; 1628 MW; 35.2% Score 31; DB 1; Length 15;  
 Query Match Best Local Similarity 62.5%; Pred No. 47;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 8 CGLYSSG 15  
 DB 7 CGRHVSCG 14  
 RESULT 3  
 TERM\_BPM2  
 ID TERM\_BPM2 STANDARD; PRT; 15 AA.  
 AC P19897;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE DNA terminal protein (protein Gp3) (Fragment).  
 GN 3 OR E.

OS Bacteriophage M2.  
 OC Viruses: dsDNA viruses, no RNA stage: Caudovirales; Podoviridae;  
 OC phi-29-like viruses.  
 OX NCBI\_TaxID=10751;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90128268; PubMed=2515115;  
 RA Matsumoto K., Takano H., Kim C.I., Hirokawa H.;  
 RT "Primary structure of bacteriophage M2 DNA polymerase: conserved  
 RT segments within protein-priming DNA polymerases and DNA polymerase I  
 RT of *Escherichia coli*.";  
 RL gene 84:247-255(1989).  
 CC -1- FUNCTION: DNA TERMINAL PROTEIN IS LINKED TO THE 5' ENDS OF  
 CC BOTH STRANDS OF THE GENOME THROUGH A PHOSPHODIESTER BOND BETWEEN  
 CC THE BETA-TERMINAL GROUP OF A SERINE RESIDUE AND THE 5'-PHOSPHATE  
 CC OF THE TERMINAL DEOXYADENYLATE. THIS PROTEIN IS ESSENTIAL FOR DNA  
 CC REPLICATION AND IS INVOLVED IN THE PRIMING OF DNA ELONGATION.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M33144; AAA32367.1; -;  
 DR PTR; PQ0017; PQ0017;  
 KW Early protein; DNA replication; DNA priming;  
 KW Covalent protein-DNA linkage.  
 FT SITE 1 1  
 FT NON\_TER 5 7  
 FT ACT\_SITE 5 7  
 SQ SEQUENCE 15 AA; 1797 MW; D3CBAFF8759DEA06 CRC64;  
 Query Match 29.5%; Score 26; DB 1; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 DRYPSGNCL 10  
 Db 1 DRYEGDVL 10  
 ||| |  
 RESULT 4  
 ID TAL3\_TREME STANDARD; PRT; 13 AA.  
 AC P01370;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1991 (Rel. 20, Last annotation update)  
 DE Tremorogen A-13.  
 DE Tremella mesenterica (Jelly fungus).  
 OS Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
 OC Tremellomycetidae; Tremellales; Tremellaceae; Tremella.  
 OX NCBI\_TaxID=5217;  
 RN [1]  
 RP SEQUENCE.  
 RA Sakagami Y., Yoshida M., Isogai A., Suzuki A.;  
 RT "Peptide sex hormones inducing conjugation tube formation in  
 RT compatible mating-type cells of *Tremella mesenterica*.";  
 RL Science 212:1525-1527(1981).  
 CC -1- FUNCTION: TREMEROGEN A-13 IS PRODUCED BY THE A MATING-TYPE CELLS  
 CC AND INDUCES FORMATION OF CONJUGATION TUBES IN A MATING-TYPE CELLS.  
 DR PR; A01641; JTUG3.  
 DR Lipoprotein; Prenylation; Pheromone.  
 KW Lipid 13 13  
 KW FARNESYL.  
 SQ SEQUENCE 13 AA; 1204 MW; 680304A9697BA864 CRC64;  
 Query Match 26.1%; Score 23; DB 1; Length 13;  
 Best Local Similarity 80.0%; Pred. No. 7.3e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 PSGNC 8

Db 9 PSGVC 13  
 ||| |  
 RESULT 5  
 ID MARI\_ALTSP STANDARD; PRT; 14 AA.  
 AC P29399;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Marinosstatin C-2 [Marinosstatin C-1; Marinosstatin D].  
 OS Alteromonas sp. (Strain B-10-31).  
 OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;  
 OC Alteromonas.  
 OX NCBI\_TaxID=29456;  
 RN [1]  
 RP SEQUENCE, AND ACTIVE SITE.  
 RX MEDLINE=92176155; PubMed=1794974;  
 RA Takano R., Inada C., Kamei K., Hara S.;  
 RT "The reactive site of marinosstatin, a proteinase inhibitor from  
 RT marine *Alteromonas* sp. B-10-31.";  
 RL J. Biochem. 110:856-858(1991).  
 CC -1- FUNCTION: INHIBITS SUBTILISIN, CHYMOTRYPSIN, AND ELASTASE, BUT  
 CC NOT TRYPsin.  
 KW Serine protease inhibitor.  
 FT PEPTIDE 1 14  
 FT PEPTIDE 3 14  
 FT PEPTIDE 4 14  
 FT MOD\_RES 1 1  
 FT ACT\_SITE 6 7  
 SQ SEQUENCE 14 AA; 1644 MW; 6E7CEE92EF32E44 CRC64;  
 Query Match 26.1%; Score 23; DB 1; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 RYPS 5  
 Db 7 RYPS 10  
 ||| |  
 RESULT 6  
 ID UC14\_MAIZE STANDARD; PRT; 15 AA.  
 AC P80620;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 258)  
 DE (Fragment).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE.  
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
 RA Fernollet J.-C., Zivy M., de Vienne D.;  
 RT "The maize two dimensional gel protein database: towards an integrated  
 RT genome analysis program.";  
 RL Theor. Appl. Genet. 93:997-1005(1996).  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 5.9, ITS MW IS: 34.6 kDa.  
 DR Maize2DPAGE; P80620; COLEOPTILE.  
 DR MaizeDB; 123944; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1564 MW; CF0BAA087DBE658 CRC64;  
 Query Match 26.1%; Score 23; DB 1; Length 15;

Best Local Similarity 57.1%; Pred. No. 8.4e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 GNCGLY 12  
I: | | |  
Db 4 GDGGAY 10

## RESULT 7

SAP\_STOVA STANDARD; PRT; 9 AA.  
AC P24047;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-MAR-1992 (Rel. 21, Last annotation update)  
DE Sperm-activating peptide (SAP).  
OS Stomopneustes variolaris (Sea urchin).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
OC Echinoida; Euechinoida; Diadematacea; Phymosomatoida; Stomechinidae;  
OC Stomopneustes.  
NCBI\_TaxID=7663;  
RN [1]

## SEQUENCE, AND DISULFIDE BOND.

RP TISSUE=Egg jelly;  
RC MEDLINE=32097763; PubMed=1756858;  
RA Yoshino K.-I., Takao T., Shimonishi Y., Suzuki N.;  
RT "Determination of the amino acid sequence of an intramolecular  
RT disulfide linkage-containing sperm-activating peptide by tandem mass  
RT spectrometry.";  
RL FEBS Lett. 294:179-182(1991).  
CC -!- FUNCTION: CAUSE STIMULATION OF SPERM RESPIRATION AND MOTILITY  
CC THROUGH INTRACELLULAR ALKALINIZATION, TRANSIENT ELEVATIONS OF  
CC CAMP, cGMP AND CLACIUM LEVELS IN SPERM CELLS, AND TRANSIENT  
CC ACTIVATION AND SUBSEQUENT INACTIVATION OF THE MEMBRANE FORM OF  
CC GUANYLATE CYCLASE.  
DR PIR: S19329; S19329. 8  
FT DISULFID 3  
SQ SEQUENCE 9 AA: 1010 MW; C469B3387B076EB9 CRC64;

Query Match 25.0%; Score 22; DB 1; Length 9;  
Best Local Similarity 60.0%; Pred. No. 1.1e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSGNC 8  
I: | | |  
Db 4 PEGKC 8

## RESULT 8

XA2\_CONGE STANDARD; PRT; 13 AA.  
AC P01520;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Alpha-conotoxin GII.  
OS Conus geographus (Geography cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
NCBI\_TaxID=6491;  
RN [1]

RP SEQUENCE.  
RX MEDLINE=81191854; PubMed=7014556;  
RA Gray W.R., Luque F.A., Olivera B.M., Barrett J., Cruz L.J.;  
RT "Peptide toxins from Conus geographus venom.";  
RL J. Biol. Chem. 256:4734-4740(1981).  
RN [2]  
RP DISULFIDE BONDS.  
RX PubMed=646616;  
RA Gray W.R., Luque F.A., Galyean R., Atherton E., Sheppard R.C.,  
RA Stone B.L., Reyes A., Alford J., McIntosh M., Olivera B.M.,  
RA Cruz L.J., Rivier J.;  
RT "Conotoxin GI: disulfide bridges, synthesis, and preparation of

RT iodinated derivatives";  
RL Biochemistry 23:2796-2802(1984).  
CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they  
CC bind to the nicotinic acetylcholine receptors (nAChR) and thus  
CC inhibit them.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE  
CC FAMILY.

DR PIR: A01783; NTKN2G.  
DR HSSP; P56973; I845.  
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;  
KW Acetylcholine receptor inhibitor; Amidation.  
FT DISULFID 2 7  
FT DISULFID 3 13  
FT MOD\_RES 13 13  
FT SEQUENCE 13 AA: 1422 MW; DEEE831C39297EBD CRC64;

Query Match 25.0%; Score 22; DB 1; Length 13;  
Best Local Similarity 50.0%; Pred. No. 1e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLYS 13  
I: | | |  
Db 7 CGKHS 12

## RESULT 9

CXAL\_CONMA STANDARD; PRT; 14 AA.  
ID CXAL\_CONMA  
AC P01521;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Alpha-conotoxin MI (MI).  
OS Conus magus (Magus cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
NCBI\_TaxID=6492;  
OX [1]  
RN SEQUENCE.  
RX MEDLINE=83073458; PubMed=7149738;  
RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;  
RT "Isolation and structure of a peptide toxin from the marine snail  
RT Conus magus.";  
RL Arch. Biochem. Biophys. 218:329-334(1982).  
RN [2]

RP DISULFIDE BONDS.  
RX MEDLINE=84032400; PubMed=6630187;  
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;  
RT "Conotoxin MI. Disulfide bonding and conformational states.";  
RL J. Biol. Chem. 258:12247-12251(1983).  
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
CC INHIBIT THEM.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE  
CC FAMILY.

DR PIR: A01784; NTKNIM.  
DR HSSP; P56973; I845.  
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;  
KW Acetylcholine receptor inhibitor; Amidation.  
FT DISULFID 3 8  
FT DISULFID 4 14  
FT MOD\_RES 14 14  
FT SEQUENCE 14 AA: 1499 MW; DEEE91898BF5E5BD CRC64;

Query Match 25.0%; Score 22; DB 1; Length 14;  
Best Local Similarity 66.7%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYS 13

```

Db      8 CGKNYS 13

RESULT 10
TRP4_LEUMA
ID TRP4_LEUMA STANDARD; PRT; 9 AA.
AC P81736;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tachykinin-related peptide 4 (LentRP 4).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE=Midgut;
MEDLINE=97053012; PubMed=8897641;
Muren J.E., Naessel D.R.;
"Isolation of five tachykinin-related peptides from the midgut of
RT the cockroach Leucophaea maderae: existence of N-terminally extended
RL isoforms.";
RL Regul. Pept. 65:185-196(1996).
CC -!- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC -!- TISSUE SPECIFICITY: MIDGUT.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW Tachykinin; Neuropeptide; Anidation.
FT MOD_RES 9 9 AMIDATION.
OX NCBI_TaxID=953;
SQ SEQUENCE 9 AA; 953 MW; 2403987699C865A7 CRC64;

Query Match 23.9%; Score 21; DB 1; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.1e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PSNGNGL 10
||| |
Db 2 PSFGMG 8

RESULT 11
C1QA_RAT
ID C1QA_RAT STANDARD; PRT; 15 AA.
AC P31720;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Complement C1q subcomponent, A chain (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC MEDLINE=93218657; PubMed=8464426;
Wing M.G., Seilly D.J., Bridgman D.J., Harrison R.A.;
"Rapid isolation and biochemical characterization of rat C1 and C1q.";
RT Mol. Immunol. 30:433-440(1993).
CC -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. THE C1Q SUBCOMPONENT IS
CC COMPOSED OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED
CC DIMERS OF THE A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-
CC LINKED DIMERS OF THE C CHAIN. IN ADDITION TO THE MAJOR A:B AND C:C

CC DIMER BANDS. RAT, UNLIKE HUMAN C1Q, CONTAINED MINOR DIMER SPECIES.
CC -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC InterPro: IPR001073; Clq.
CC PROSITE: PS01113; Clq; PARTIAL.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat.
FT DISULFID 4 4 INTERCHAIN (WITH C-4 IN B CHAIN).
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1488 MW; 1B3D8000B7793965 CRC64;

Query Match 23.9%; Score 21; DB 1; Length 15;
Best Local Similarity 44.4%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 RYPGNCGL 10
| | | |
Db 5 RAPNGKGV 13

RESULT 12
FGF1_CANFA
ID FGF1_CANFA STANDARD; PRT; 15 AA.
AC P18651;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Heparin-binding growth factor 1 (HBGF-1) (Acidic fibroblast growth
DE factor) (AFGF) (Alpha-endothelial cell growth factor) (fragment).
GN FGF1 OR FGF-1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Flissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RC MEDLINE=892331704; PubMed=2714282;
Quinkler W., Maasberg M., Bernotat-Danielowski S., Luethe N.,
RA Sharma H.S., Schaper W.;
"Isolation of heparin-binding growth factors from bovine, porcine and
RT canine hearts.";
RL Eur. J. Biochem. 181:67-73(1989).
CC -!- FUNCTION: THE HEPARIN-BINDING GROWTH FACTORS ARE ANGIOGENIC AGENTS
CC IN VIVO AND ARE POTENT MITOGENS FOR A VARIETY OF CELL TYPES IN
CC VITRO. THERE ARE DIFFERENCES IN THE TISSUE DISTRIBUTION AND
CC CONCENTRATION OF THESE 2 GROWTH FACTORS.
CC -!- SUBUNIT: MONOMER.
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS HEPARIN, ALTHOUGH LESS STRONGLY
CC THAN DOES BFGF.
CC -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
DR InterPro: IPR002209; HB/F_growthfact.
DR PIR: S03955; S03955.
KW Growth factor; Mitogen; Angiogenesis; Heparin-binding.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1732 MW; 53CC9A3CADDADAAL CRC64;

Query Match 23.9%; Score 21; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 10 LYSSSG 15
|| | |
Db 8 LYXSNG 13

RESULT 13
PPCK_FASHE
ID PPCK_FASHE STANDARD; PRT; 10 AA.
AC P80525;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32)

```

```

DE (Phosphoenolpyruvate carboxylase) (PEPCK) (Newly excysted juvenile
DE protein 1) (Fragment).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fasciolidae; Fasciolidae; Fasciola.
ON NCBI_TaxID=6192;
RX SEQUENCE.
RX MEDLINE=95366993; PubMed=7639732;
RA Tkalcovic J., Ashman K., Meusen E.;
RT "Fasciola hepatica: rapid identification of newly excysted juvenile
RT proteins."
RL Biochem. Biophys. Res. Commun. 213:169-174(1995).
CC -!- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -!- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]
CC FAMILY.
DR InterPro: IPR000364; PEP_carboxykin.
DR PROSITE: PS00505; PEPCK_GTP; PARTIAL.
KW Lyase; Decarboxylase; GTP-binding.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1069 MW; 8393AG187AA9C87A CRC64;

Query Match 22.7%; Score 20; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 YPSG 6
DB 1 YPDG 4

RESULT 14
TKL1 LOCMI STANDARD; PRT; 9 AA.
AC PL6223;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Locustatachykinin I (TK-I).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT "Locustatachykinin I and II, two novel insect neuropeptides with
RT homology to peptides of the vertebrate tachykinin family."
RL FEBS Lett. 261:397-401(1990).
CC -!- FUNCTION: MYOACTIVE PEPTIDE. STIMULATES THE CONTRACTION OF THE
CC OVIDUCT AND FOREGUT.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
DR PIR: S08265; ECLQ1M.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 9 9
SQ SEQUENCE 9 AA; 939 MW; 2389C86B59C865A7 CRC64;

Query Match 21.6%; Score 19; DB 1; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.1e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSNGCGL 10
DB 2 PSFGYGV 8

RESULT 15
HTF_HELZE STANDARD; PRT; 10 AA.
ID HTF_HELZE

```

```

AC PL6353;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Hypertrehalosemic hormone (Hez-HRTH).
OS Heliothis zea (Corn earworm) (Bollworm).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Diptera; Noctuidae; Noctuidae; Heliothinae; Helicoverpa.
OX NCBI_TaxID=7113;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=88326324; PubMed=3415690;
RA Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Bird T.G.,
RA Tseng C.M., Zhang Y.S., Hayes D.K.;
RT "Isolation and primary structure of a neuro peptide hormone from
RT Heliothis zea with hypertrehalosemic and adipokinetic activities."
RL Biochem. Biophys. Res. Commun. 155:344-350(1988).
CC -!- FUNCTION: HYPERTREHALOSEMIC FACTORS ARE NEUROPEPTIDES THAT
CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR: A31571; A31571.
DR InterPro: IPR002047; AKH.
DR PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1096 MW; 8E70367865A5B9D1 CRC64;

Query Match 21.6%; Score 19; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.4e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 LYVSSG 15
DB 2 LTFSSG 7

Search completed: December 2, 2002, 07:02:02
Job time : 13 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:59:26 ; Search time 28 Seconds  
(without alignments)  
110.382 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRYPSGNGGLYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 3161

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_plage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 27    | 30.7        | 10     | 10    | P82438 nicotiana t |
| 2          | 25    | 28.4        | 13     | 5     | Q8T6E9 drosophila  |
| 3          | 25    | 28.4        | 13     | 5     | Q8STI5 drosophila  |
| 4          | 24    | 27.3        | 9      | 4     | Q96T78 drosophila  |
| 5          | 24    | 27.3        | 11     | 2     | Q9AIX6 homo sapien |
| 6          | 24    | 27.3        | 11     | 5     | Q9NL65 carsonella  |
| 7          | 22    | 25.0        | 10     | 6     | Q9WP04 ateleis suu |
| 8          | 22    | 25.0        | 11     | 13    | Q90WA2 gallus gall |
| 9          | 22    | 25.0        | 15     | 4     | Q00604 homo sapien |
| 10         | 22    | 25.0        | 15     | 8     | Q95770 cyclura ric |
| 11         | 22    | 25.0        | 15     | 13    | Q9PRZ9 microgogoni |
| 12         | 21    | 23.9        | 10     | 13    | Q90Y93 gallus gall |
| 13         | 21    | 23.9        | 12     | 10    | Q945C4 crypthecodi |
| 14         | 21    | 23.9        | 13     | 4     | Q9UPE7 homo sapien |
| 15         | 21    | 23.9        | 15     | 10    | Q9S8V7 triticum ae |
| 16         | 20    | 22.7        | 12     | 2     | Q93U04 escherichia |

|    |    |      |    |    |                     |
|----|----|------|----|----|---------------------|
| 17 | 20 | 22.7 | 12 | 4  | Q9UMQ9 homo sapien  |
| 18 | 20 | 22.7 | 13 | 11 | Q9WTZ6 mus musculus |
| 19 | 20 | 22.7 | 13 | 11 | P82808 rattus norv  |
| 20 | 20 | 22.7 | 14 | 2  | Q45876 clostridium  |
| 21 | 20 | 22.7 | 14 | 2  | Q45872 clostridium  |
| 22 | 20 | 22.7 | 14 | 12 | Q9PY99 murine hepa  |
| 23 | 20 | 22.7 | 15 | 4  | Q9UJ51 homo sapien  |
| 24 | 20 | 22.7 | 15 | 12 | Q86865 lymphocytic  |
| 25 | 20 | 22.7 | 15 | 12 | Q86867 lymphocytic  |
| 26 | 20 | 22.7 | 15 | 12 | Q86869 lymphocytic  |
| 27 | 19 | 21.6 | 7  | 10 | Q9C5B3 arabidopsis  |
| 28 | 19 | 21.6 | 8  | 7  | Q95213 oryctolagus  |
| 29 | 19 | 21.6 | 10 | 4  | Q9UCR0 homo sapien  |
| 30 | 19 | 21.6 | 10 | 4  | Q15342 homo sapien  |
| 31 | 19 | 21.6 | 13 | 4  | Q9UC27 homo sapien  |
| 32 | 19 | 21.6 | 15 | 1  | Q9UWH6 thermococcu  |
| 33 | 19 | 21.6 | 15 | 10 | Q9S8L4              |
| 34 | 19 | 21.6 | 15 | 11 | Q9QVA7              |
| 35 | 18 | 20.5 | 9  | 6  | Q9TRU7 bos taurus   |
| 36 | 18 | 20.5 | 9  | 11 | C08979 mus musculu  |
| 37 | 18 | 20.5 | 9  | 11 | Q9QWT0 mus musculu  |
| 38 | 18 | 20.5 | 10 | 2  | Q9X533 escherichia  |
| 39 | 18 | 20.5 | 10 | 2  | Q9X534 leclercia a  |
| 40 | 18 | 20.5 | 10 | 2  | Q48469 klebsiella   |
| 41 | 18 | 20.5 | 11 | 5  | Q9TWM2 aplysia cal  |
| 42 | 18 | 20.5 | 11 | 12 | Q86864 lymphocytic  |
| 43 | 18 | 20.5 | 11 | 12 | Q86866 lymphocytic  |
| 44 | 18 | 20.5 | 11 | 12 | Q86868 lymphocytic  |
| 45 | 18 | 20.5 | 12 | 8  | Q95EL4 dendrochilu  |

## ALIGNMENTS

## RESULT 1

P82438 PRELIMINARY; PRT; 10 AA.  
AC P82438;  
DT 01-JUN-2000 (TREMBLrel. 14, Created)  
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE 50 kDa cell wall protein (Fragment).  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4097;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=CV. PETIT HAVANA;  
RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,  
RA Wojtaszek P., Bolwell G.P.;  
RT "Proteomic study of secondary cell wall proteins from transformed  
RT tobacco culture."  
RL Planta 0:0-0(2000).  
CC -1- SUBCELLULAR LOCATION: CELL WALL.  
CC -1- TISSUE SPECIFICITY: XYLEM.  
KW Cell wall.  
FT NON\_TER.  
SQ SEQUENCE 10 AA; 1126 MW; C68E32486AF77B46 CRC64;

Query Match 30.7%; Score 27; DB 10; Length 10;  
Best Local Similarity 66.7%; Pred. No. 4.2e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RYPSGN 7

Db 3 QYPSGN 8

## RESULT 2

Q8T6E9 PRELIMINARY; PRT; 13 AA.  
ID Q8T6E9

AC Q8T6F9;  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
DE Myocyte enhancing factor 2 (Fragment).  
GN ME2.  
OS Drosophila miranda (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7237;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP Schaeffer S.W., Goetting-Minesky M.P., Kovacevic M., Peoples J.,  
RA Graybill J.L., Miller J.M., Kim K., Nelson J.G., Anderson W.W.;  
RT "Evolutionary genomics of inversions in Drosophila pseudoobscura:  
RT Modes of selection."  
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF476617; AAL91815.1; -  
DR EMBL; AF476617; AAL91815.1; -  
FT NON\_TER 1  
FT 13  
SQ SEQUENCE 13 AA; 1313 MW; 15BF380B6BE05050 CRC64;  
Query Match 28.4%; Score 25; DB 5; Length 13;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 YPSGN 7  
DB 8 YPSGS 12  
RESULT 3  
Q8STI5 PRELIMINARY; PRT; 13 AA.  
AC Q8STI5;  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
DE Myocyte enhancing factor 2 (Fragment).  
GN ME2.  
OS Drosophila pseudoobscura (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7237;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP Schaeffer S.W., Goetting-Minesky M.P., Kovacevic M., Peoples J.,  
RA Graybill J.L., Miller J.M., Kim K., Nelson J.G., Anderson W.W.;  
RT "Evolutionary genomics of inversions in Drosophila pseudoobscura:  
RT Modes of selection."  
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF476728; AAL91726.1; -  
DR EMBL; AF476729; AAL91727.1; -  
DR EMBL; AF476730; AAL91728.1; -  
DR EMBL; AF476731; AAL91729.1; -  
DR EMBL; AF476732; AAL91730.1; -  
DR EMBL; AF476733; AAL91731.1; -  
DR EMBL; AF476734; AAL91732.1; -  
DR EMBL; AF476735; AAL91733.1; -  
DR EMBL; AF476736; AAL91734.1; -  
DR EMBL; AF476737; AAL91735.1; -  
DR EMBL; AF476738; AAL91736.1; -  
DR EMBL; AF476739; AAL91737.1; -  
DR EMBL; AF476740; AAL91738.1; -  
DR EMBL; AF476741; AAL91739.1; -  
DR EMBL; AF476742; AAL91740.1; -  
DR EMBL; AF476743; AAL91741.1; -  
DR EMBL; AF476744; AAL91742.1; -  
DR EMBL; AF476745; AAL91743.1; -  
DR EMBL; AF476746; AAL91744.1; -  
DR EMBL; AF476747; AAL91745.1; -  
DR EMBL; AF476748; AAL91746.1; -  
DR EMBL; AF476749; AAL91747.1; -  
DR EMBL; AF476750; AAL91748.1; -  
DR EMBL; AF476751; AAL91749.1; -  
DR EMBL; AF476752; AAL91750.1; -  
DR EMBL; AF476753; AAL91751.1; -  
DR EMBL; AF476754; AAL91752.1; -  
DR EMBL; AF476755; AAL91753.1; -  
DR EMBL; AF476756; AAL91754.1; -  
DR EMBL; AF476757; AAL91755.1; -  
DR EMBL; AF476758; AAL91756.1; -  
DR EMBL; AF476759; AAL91757.1; -  
DR EMBL; AF476760; AAL91758.1; -  
DR EMBL; AF476761; AAL91759.1; -  
DR EMBL; AF476762; AAL91760.1; -  
DR EMBL; AF476763; AAL91761.1; -  
DR EMBL; AF476764; AAL91762.1; -  
DR EMBL; AF476765; AAL91763.1; -  
DR EMBL; AF476766; AAL91764.1; -  
DR EMBL; AF476767; AAL91765.1; -  
DR EMBL; AF476768; AAL91766.1; -  
DR EMBL; AF476769; AAL91767.1; -  
DR EMBL; AF476770; AAL91768.1; -  
DR EMBL; AF476771; AAL91769.1; -  
DR EMBL; AF476772; AAL91770.1; -  
DR EMBL; AF476773; AAL91771.1; -  
DR EMBL; AF476774; AAL91772.1; -  
DR EMBL; AF476775; AAL91773.1; -  
DR EMBL; AF476776; AAL91774.1; -  
DR EMBL; AF476777; AAL91775.1; -  
DR EMBL; AF476778; AAL91776.1; -  
DR EMBL; AF476779; AAL91777.1; -  
DR EMBL; AF476780; AAL91778.1; -  
DR EMBL; AF476781; AAL91779.1; -  
DR EMBL; AF476782; AAL91780.1; -  
DR EMBL; AF476783; AAL91781.1; -  
DR EMBL; AF476784; AAL91782.1; -  
DR EMBL; AF476785; AAL91783.1; -  
DR EMBL; AF476786; AAL91784.1; -  
DR EMBL; AF476787; AAL91785.1; -  
DR EMBL; AF476788; AAL91786.1; -  
DR EMBL; AF476789; AAL91787.1; -  
DR EMBL; AF476790; AAL91788.1; -  
DR EMBL; AF476791; AAL91789.1; -  
DR EMBL; AF476792; AAL91790.1; -  
DR EMBL; AF476793; AAL91791.1; -  
DR EMBL; AF476794; AAL91792.1; -  
DR EMBL; AF476795; AAL91793.1; -  
DR EMBL; AF476796; AAL91794.1; -  
DR EMBL; AF476797; AAL91795.1; -  
DR EMBL; AF476798; AAL91796.1; -  
DR EMBL; AF476799; AAL91797.1; -  
DR EMBL; AF476800; AAL91798.1; -  
DR EMBL; AF476801; AAL91799.1; -  
DR EMBL; AF476802; AAL91800.1; -  
DR EMBL; AF476803; AAL91801.1; -  
DR EMBL; AF476804; AAL91802.1; -  
DR EMBL; AF476805; AAL91803.1; -  
DR EMBL; AF476806; AAL91804.1; -  
DR EMBL; AF476807; AAL91805.1; -  
DR EMBL; AF476808; AAL91806.1; -  
DR EMBL; AF476809; AAL91807.1; -  
DR EMBL; AF476810; AAL91808.1; -  
DR EMBL; AF476811; AAL91809.1; -  
DR EMBL; AF476812; AAL91810.1; -  
DR EMBL; AF476813; AAL91811.1; -  
DR EMBL; AF476814; AAL91812.1; -  
DR EMBL; AF476815; AAL91813.1; -  
DR EMBL; AF476816; AAL91814.1; -  
FT NON\_TER 1  
FT 13  
SQ SEQUENCE 13 AA; 1313 MW; 15BF380B6BE05050 CRC64;

Query Match 28.4%; Score 25; DB 5; Length 13;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPSGN 7  
IIII:  
Db 8 YPSGS 12

## RESULT 4

Q96T78 ID Q96T78 PRELIMINARY; PRT; 9 AA.  
AC Q96T78  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Contactin-associated protein 2 (Fragment).  
GN CNTNAP2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OC NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21250995; PubMed=11352571;  
RA Nakabayashi K., Scherer S.W.;  
RT "The human contactin-associated protein 2 (CNTNAP2) spans over 2 Mb of  
RT DNA at chromosome 7q35.";  
RL Genomics 73:108-112(2001).  
DR EMBL; AF318225; AK4906.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 1080 MW; 9139A2D5A77B51EA CRC64;

Query Match 27.3%; Score 24; DB 4; Length 9;  
Best Local Similarity 80.0%; Pred. No. 6.7e+05;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DRYPS 5  
IIII  
Db 2 DRYPS 6

## RESULT 5

Q9AIY6 ID Q9AIY6 PRELIMINARY; PRT; 11 AA.  
AC Q9AIY6  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (Fragment).  
GN TRPS.  
OS Carsonella ruddii.  
OC Bacteria; Proteobacteria; gamma subdlvision; Candidatus Carsonella.  
OC NCBI\_TaxID=114186;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20336438; PubMed=10877784;  
RA Thao M.L., Moran N.A., Abbot P., Brennan E.B., Burchhardt D.H.,  
RA Baumann P.;  
RT "Cospeciation of psyllids and their primary prokaryotic  
RT endosymbionts.";  
RL Appl. Environ. Microbiol. 66:2898-2905(2000).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21125546; PubMed=11222582;  
RA Clark M.A., Baumann L., Thao M.L., Moran N.A., Baumann P.;  
RT "Degenerative Minimalism in the Genome of a Psyllid Endosymbiont.";  
RL J. Bacteriol. 183:1853-1861(2001).  
DR EMBL; AF211138; AK45388.1; -.  
KW Aminoacyl-tRNA synthetase.  
FT NON\_TER 1 1  
SQ SEQUENCE 11 AA; 1295 MW; 0CA993A5345B5720 CRC64;

Query Match 27.3%; Score 24; DB 2; Length 11;  
Best Local Similarity 42.9%; Pred. No. 1.5e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 NCGLYS 13  
IIII:  
Db 1 NGSILYN 7

## RESULT 6

Q9NL65 ID Q9NL65 PRELIMINARY; PRT; 11 AA.  
AC Q9NL65  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE ASABF-Delta (Fragment).  
GN ASABF-Delta.  
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascarididae; Ascaridoidea;  
OC Ascarididae; Ascaris.  
OC NCBI\_TaxID=6253;  
[1]  
RP SEQUENCE FROM N.A.  
RX Kato Y.;  
RT "Ascaris suum asabf-delta gene, exon 2.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB029815; BAA89496.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1187 MW; 8BADD0CD1EAB5861 CRC64;

Query Match 27.3%; Score 24; DB 5; Length 11;  
Best Local Similarity 66.7%; Pred. No. 1.5e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 NCGLYY 12  
IIII  
Db 1 NCGTGY 6

## RESULT 7

Q8WP04 ID Q8WP04 PRELIMINARY; PRT; 10 AA.  
AC Q8WP04  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Oculocutaneous albinism type II (Fragment).  
GN OCA2.  
OS Ateles belzebuth chamek (Chamek spider monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.  
OC NCBI\_TaxID=118643;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21636947; PubMed=11778686;  
RA Seunavez H.N., Lima C.R., Lemos B., Bonvicino C.R., Moreira M.A.M.,  
RA Canavez F.C.;  
RT "Gene assignment in Ateles paniscus chamek (Platyrrhini, Primates).  
RT Allocation of 18 markers of human syntenic groups 1.2,7,14,15,17 and  
RT 22.";  
RL Chromosome Res. 9:631-639(2001).  
DR EMBL; AF375652; RAL31489.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 10 AA; 901 MW; 22DF477DD87EA5B8 CRC64;

Query Match 25.0%; Score 22; DB 6; Length 10;  
Best Local Similarity 75.0%; Pred. No. 2.9e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 GNCG 9

Db 3 GSCG 6

RESULT 8  
Q90WA2  
ID Q90WA2 PRELIMINARY; PRT; 11 AA.  
AC Q90WA2;  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE Extracellular fatty acid binding protein (fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
TX TISSUE=BLOOD;  
WA Wang Q., Li N., Deng X., Li H.;  
RT "Single Nucleotide Polymorphism Analysis on Encoding Region of  
RT Extracellular Fatty Acid Binding Protein Genes and Their Associations  
RT With the Fatness Trait in Chicken.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF402001; AAK94082.1; -  
FT NON\_TER 1 1  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1277 MW; 873417F02B187AAA CRC64;

Query Match 25.0%; Score 22; DB 13; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 LYYS 13  
Db 8 LYYS 11

RESULT 9  
O00604  
ID O00604 PRELIMINARY; PRT; 15 AA.  
AC O00604;  
DT 01-JUL-1997 (TRENBLrel. 04, Created)  
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE Type I collagen alpha 1(I) chain (fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96336021; PubMed=8757037;  
RA Mayer S.A., Rubin B.S., Starman B.J., Byers P.H.;  
RT "Spontaneous multivessel cervical artery dissection in a patient with  
RT a substitution of alanine for glycine (G13A) in the alpha 1 (I) chain  
RT of type I collagen.";  
RL Neurology 47:552-556(1996).  
DR EMBL: S83315; AAB50780.2; -  
FT NON\_TER 1 1  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1413 MW; 4CA8D720AAF780FC CRC64;

Query Match 25.0%; Score 22; DB 4; Length 15;  
Best Local Similarity 71.4%; Pred. No. 4.4e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 FSGNGL 10  
Db 2 PSGRGL 8

RESULT 10

Q95770  
ID Q95770 PRELIMINARY; PRT; 15 AA.  
AC Q95770;  
DT 01-FEB-1997 (TRENBLrel. 02, Created)  
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE NADH dehydrogenase subunit 4 (fragment).  
GN N04.  
OS Cyclura ricordi.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Iguania; Iguanidae; Iguaninae; Cyclura.  
OX NCBI\_TaxID=51215;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97019047; PubMed=8865663;  
RA Sites J.W. Jr., Davis S.K., Guerra T., Iverson J.B., Snell H.L.;  
RT "Character congruence and phylogenetic signal in molecular and  
RT morphological data sets: a case study in the living Iguanas (Squamata,  
RT Iguanidae).";  
RL Mol. Biol. Evol. 13:1087-1105(1996).  
DR EMBL: U66237; AAB07473.1; -  
KW Mitochondrion.  
FT NON\_TER 1 1  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1715 MW; 8327178E7927A57E CRC64;

Query Match 25.0%; Score 22; DB 8; Length 15;  
Best Local Similarity 50.0%; Pred. No. 4.4e+03;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 7 NCGLYYSS 14  
Db 2 NCRHSTS 9

RESULT 11  
Q9PRZ9  
ID Q9PRZ9 PRELIMINARY; PRT; 15 AA.  
AC Q9PRZ9;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
DE Gonadotropin I beta subunit (fragment).  
OS Micropogonias undulatus (Atlantic croaker).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
OC Sciaenidae; Micropogonias.  
OX NCBI\_TaxID=29154;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94010173; PubMed=8405898;  
RA Copeland P.A., Thomas P.;  
RT "Isolation of gonadotropin subunits and evidence for two distinct  
RT gonadotropins in Atlantic croaker (Micropogonias undulatus).";  
RL Gen. Comp. Endocrinol. 91:115-125(1993).  
SQ SEQUENCE 15 AA; 1732 MW; 81709992C3D86A4A CRC64;

Query Match 25.0%; Score 22; DB 13; Length 15;  
Best Local Similarity 75.0%; Pred. No. 4.4e+03;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 8 CGLV 11  
Db 4 CGFY 7

RESULT 12  
Q90Y93  
ID Q90Y93 PRELIMINARY; PRT; 10 AA.

```

AC Q90Y93;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Growth hormone (Fragment).
GN GH.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Kansaku N., Nakada A., Yagi E., Okabayashi H., Guemene D.;
RT "Genetic variation of chicken growth hormone gene.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB061722; BAB69037.1; -
FT NON_TER 1
FT NON_TER 10
FT NON_TER 10
SQ SEQUENCE 10 AA; 1155 MW; 684175175A40AAB CRC64;
Query Match 23.9%; Score 21; DB 13; Length 10;
Best Local Similarity 66.7%; Pred. No. 4.3e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 DRYPSG 6
DB 4 DRSRPG 9

RESULT 13
Q945C4
ID Q945C4 PRELIMINARY; PRT; 12 AA.
AC Q945C4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Beta-tubulin (Fragment).
OS Cryptocodium cohni (Dinoflagellate).
OC Eukaryota; Alveolata; Dinophyceae; Gonyaulacales; Cryptophyceae;
OC Cryptocodium.
OX NCBI_TaxID=2866;
RN [1]
RP SEQUENCE FROM N.A.
RA Guillebaud D., Derelle E., Lozano J.C., Bingham S., Moreau H.;
RT "A single TBP-like protein is present in the marine unicellular
RT organism: the dinoflagellate Cryptocodium cohni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF417567; AAL15905.1; -
FT NON_TER 12
FT NON_TER 12
SQ SEQUENCE 12 AA; 1299 MW; F2AE9D1D4166D051 CRC64;
Query Match 23.9%; Score 21; DB 10; Length 12;
Best Local Similarity 75.0%; Pred. No. 5.2e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 GNCG 9
DB 9 GQCG 12

RESULT 14
Q9UPE7
ID Q9UPE7 PRELIMINARY; PRT; 13 AA.
AC Q9UPE7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Inosine monophosphatase 2 (Fragment).
GN IMPA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoshikawa T., Turner G., Esterling L.E., Sanders A.R.,
RT Detera-Wadleigh S.D.;
RT "A novel human myo-inositol monophosphatase gene, IMP.18p, maps to a
RT susceptibility region for bipolar disorder.";
RL Mol. Psychiatry 2:393-397(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Yoshikawa T., Padigar M., Karkera J.D., Sharma M., Berrettini W.H.,
RT Esterling L.E., Detera-Wadleigh S.D.;
RT "Genomic structure and novel variants of myo-inositol monophosphatase
RT 2.";
RL Mol. Psychiatry 5:165-171(2000).
DR EMBL; AF025882; AAD22136.1; -
DR EMBL; AF025881; AAD22136.1; JOINED.
FT NON_TER 1
FT NON_TER 13
FT NON_TER 13
SQ SEQUENCE 13 AA; 1589 MW; F3415D841F48D401 CRC64;
Query Match 23.9%; Score 21; DB 4; Length 13;
Best Local Similarity 60.0%; Pred. No. 5.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 DRYPS 5
DB 1 ENFPS 5

RESULT 15
Q9SV7
ID Q9SV7 PRELIMINARY; PRT; 15 AA.
AC Q9SV7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE High-molecular-weight glutenin subunit (Fragment).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE.
RA Tao H.P., Adalsteins A.E., Kasarda D.D.;
RT "Intermolecular disulfide bonds link specific high-molecular-weight
RT glutenin subunits in wheat endosperm.";
RL Biochim. Biophys. Acta 1159:13-21(1992).
SQ SEQUENCE 15 AA; 1559 MW; CDIA6F573C945AFD CRC64;
Query Match 23.9%; Score 21; DB 10; Length 15;
Best Local Similarity 50.0%; Pred. No. 6.6e+03;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 PSNGGLY 11
DB 4 PQGXGY 11

Search completed: December 2, 2002, 07:02:38
Job time : 31 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:54:10 ; Search time 91 Seconds  
(without alignments)  
21.964 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRPSNGCLYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 247533

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 88    | 100.0       | 15     | 20    | AAW8237     |
| 2          | 33    | 37.5        | 9      | 22    | AAU23853    |
| 3          | 33    | 37.5        | 9      | 22    | AAU24319    |
| 4          | 33    | 37.5        | 9      | 22    | AAU24423    |
| 5          | 33    | 37.5        | 9      | 22    | AAU24441    |
| 6          | 33    | 37.5        | 10     | 22    | AAU24484    |
| 7          | 33    | 37.5        | 12     | 21    | AAU26044    |
| 8          | 33    | 37.5        | 12     | 22    | AAU16769    |
| 9          | 33    | 37.5        | 12     | 23    | ABJ00390    |
| 10         | 33    | 37.5        | 15     | 18    | AAW23509    |

|    |    |      |    |    |          |                    |
|----|----|------|----|----|----------|--------------------|
| 11 | 33 | 37.5 | 15 | 19 | AAW57788 | Fatty acid amide h |
| 12 | 31 | 35.2 | 10 | 22 | AAG96214 | Human complementar |
| 13 | 31 | 35.2 | 15 | 5  | AAP40328 | Sequence of conoto |
| 14 | 31 | 35.2 | 15 | 16 | AAW75265 | Alpha-conotoxin GI |
| 15 | 31 | 35.2 | 15 | 18 | AAW12727 | Physarum polyceph  |
| 16 | 30 | 34.1 | 7  | 20 | AAV30230 | Human complementa  |
| 17 | 30 | 34.1 | 10 | 22 | AAG94296 | Human complementa  |
| 18 | 30 | 34.1 | 12 | 20 | AAV29882 | Humicola lanuginos |
| 19 | 30 | 34.1 | 13 | 15 | AAW46829 | pH 2.5 acid phosph |
| 20 | 30 | 34.1 | 14 | 15 | AAW58323 | Hypotensive polype |
| 21 | 30 | 34.1 | 14 | 19 | AAW59655 | Amino acid sequenc |
| 22 | 30 | 34.1 | 14 | 22 | AAW68153 | Peptide derived fr |
| 23 | 30 | 34.1 | 14 | 22 | AAW68154 | Peptide derived fr |
| 24 | 30 | 34.1 | 14 | 22 | AAW68187 | Peptide derived fr |
| 25 | 30 | 34.1 | 14 | 22 | AAW68188 | Peptide derived fr |
| 26 | 30 | 34.1 | 15 | 16 | AAW63873 | Cryoglobulinemia t |
| 27 | 30 | 34.1 | 15 | 19 | AAW54431 | Human PS112 protei |
| 28 | 30 | 34.1 | 15 | 20 | AAV30227 | Physarum polyceph  |
| 29 | 30 | 34.1 | 15 | 21 | AAW08421 | Antigen peptide de |
| 30 | 30 | 34.1 | 15 | 22 | AAG78001 | Human actin 14 pep |
| 31 | 29 | 33.0 | 10 | 22 | AAU24373 | Human MHC class I  |
| 32 | 29 | 33.0 | 10 | 22 | AAU24478 | Human MHC molecule |
| 33 | 29 | 33.0 | 14 | 14 | AAW69320 | Gp IIB/IIIA recept |
| 34 | 29 | 33.0 | 14 | 21 | AAW11799 | Human ERV2 protein |
| 35 | 29 | 33.0 | 14 | 21 | AAV95520 | Fibrin polymerizat |
| 36 | 29 | 33.0 | 15 | 20 | AAV07427 | HSV-1 TK amino aci |
| 37 | 29 | 33.0 | 15 | 20 | AAW97159 | Transforming growt |
| 38 | 29 | 33.0 | 15 | 21 | AAV93038 | Peptide which is n |
| 39 | 28 | 31.8 | 7  | 23 | AAU11469 | Human leucocyte an |
| 40 | 28 | 31.8 | 8  | 19 | AAW63298 | Saccharomyces cere |
| 41 | 28 | 31.8 | 9  | 17 | AAW49556 | Miscellaneous pept |
| 42 | 28 | 31.8 | 10 | 22 | AAW65409 | Sperm activating p |
| 43 | 28 | 31.8 | 11 | 23 | AAW92402 | Interleukin-1 type |
| 44 | 28 | 31.8 | 11 | 23 | AAU11482 |                    |
| 45 | 28 | 31.8 | 12 | 18 | AAW15982 |                    |

#### ALIGNMENTS

RESULT 1  
AAW8237  
ID AAW8237 standard; Peptide; 15 AA.  
XX  
AC AAW8237;  
XX  
DT 15-MAR-1999 (first entry)  
XX  
DE Human prothrombinase Fg12 epitope.  
XX  
KW Prothrombinase; hf912; Fg12; human; immune coagulation; antibody;  
KW inhibitor; infection; graft rejection; glomerulonephritis; cancer;  
KW gastrointestinal disease; foetal loss; therapy; vaccine; epitope.  
XX  
OS Homo sapiens.  
XX  
PN WO9851335-A1.  
XX  
PD 19-NOV-1998.  
XX  
PF 15-MAY-1998; 98WO-CA00475.  
XX  
PR 10-OCT-1997; 97US-0061684.  
PR 15-MAY-1997; 97US-0046537.  
(LEVY/) LEVY G.  
PI  
Levy G;  
WPI; 1999-059687/05.  
DR  
Modulating immune coagulation - by using Fg12 antibodies and  
PT compounds, used to treat conditions including graft rejection and

PT foetal loss

PS Claim 4; Page 72; 105pp; English.

XX

CC This peptide corresponds to amino acid residues 364-378 of human prothrombinase Fg12 (see AA088235). A claimed method of preventing or treating a condition requiring a reduction in immune coagulation comprises administering an inhibitor of Fg12. The inhibitor is preferably an antibody that binds to the Fg12 epitope. The condition to be treated in graft rejection of foetal loss (claimed).

XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 88; DB 20; Length 15;

Best Local Similarity 100.0%; Pred. No. 1e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLYSSG 15

IIIIIIIIIIIIIIIIII

DB 1 DRYPSGNCGLYSSG 15

RESULT 2

AAU23853

ID AAU23853 standard; Peptide; 9 AA.

XX

AC AAU23853;

XX

DT 17-DEC-2001 (first entry)

XX

DE Human MHC class I molecule HLA-A1 binding 103P2D6 peptide #38.

XX

KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue; tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic; gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder; single chain monoclonal antibody; cervix; human.

XX

OS Homo sapiens.

XX

PN WO200162925-A2.

XX

PD 30-AUG-2001.

XX

PF 26-FEB-2001; 2001WO-US05996.

XX

PR 24-FEB-2000; 2000US-0184558.

XX

PT 13-JUL-2000; 2000US-0218856.

XX

PA (UROG-) UROGENESYS INC.

XX

PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;

XX

PT Challita-eid PM, Faris M, Jakobovits A;

XX

WPI; 2001-557705/62.

XX

PT New polynucleotide for treating and diagnosing prostate cancer is the

XX

PT 103P2D6 gene which encodes for 103P2D6-related proteins

XX

PS Example 15; Page 81; 132pp; English.

XX

CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and peptide fragments of the polypeptide. 103P2D6 is not expressed in normal adult tissue but is aberrantly expressed in some foetal tissues and many cancers including tumours of the prostate, testis, bladder, bone, cervix, ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its related protein and also peptide fragments of the protein are therefore useful for diagnosing and treating cancer. A vector comprising a polynucleotide which encodes a single chain monoclonal antibody, that immunospecifically binds to an 103P2D6-related protein, and a ribozyme capable of cleaving a polynucleotide having the 103P2D6 coding sequence, are both useful in the preparation of a composition for treating a patient with a cancer that expresses 103P2D6. The sequences can be used

CC in diagnostic methods to monitor the level of 103P2D6 gene products in serum, blood, urine and tissue and to thereby detect the presence of cancerous cells.

XX

SQ Sequence 9 AA;

Query Match 37.5%; Score 33; DB 22; Length 9;

Best Local Similarity 75.0%; Pred. No. 7.8e+05;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLY 12

IIIIII

DB 2 SGRGGLY 9

RESULT 3

AAU24319

ID AAU24319 standard; Peptide; 9 AA.

XX

AC AAU24319;

XX

DT 17-DEC-2001 (first entry)

XX

DE Human MHC class I molecule HLA-B7 binding 103P2D6 peptide #4.

XX

KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue; tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic; gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder; single chain monoclonal antibody; cervix; human.

XX

OS Homo sapiens.

XX

PN WO200162925-A2.

XX

PD 30-AUG-2001.

XX

PF 26-FEB-2001; 2001WO-US05996.

XX

PR 24-FEB-2000; 2000US-0184558.

XX

PT 13-JUL-2000; 2000US-0218856.

XX

PA (UROG-) UROGENESYS INC.

XX

PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;

XX

PT Challita-eid PM, Faris M, Jakobovits A;

XX

WPI; 2001-557705/62.

XX

PT New polynucleotide for treating and diagnosing prostate cancer is the

XX

PT 103P2D6 gene which encodes for 103P2D6-related proteins

XX

PS Example 15; Page 94; 132pp; English.

XX

CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and peptide fragments of the polypeptide. 103P2D6 is not expressed in normal adult tissue but is aberrantly expressed in some foetal tissues and many cancers including tumours of the prostate, testis, bladder, bone, cervix, ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its related protein and also peptide fragments of the protein are therefore useful for diagnosing and treating cancer. A vector comprising a polynucleotide which encodes a single chain monoclonal antibody, that immunospecifically binds to an 103P2D6-related protein, and a ribozyme capable of cleaving a polynucleotide having the 103P2D6 coding sequence, are both useful in the preparation of a composition for treating a patient with a cancer that expresses 103P2D6. The sequences can be used in diagnostic methods to monitor the level of 103P2D6 gene products in serum, blood, urine and tissue and to thereby detect the presence of cancerous cells.

XX

SQ Sequence 9 AA;

Query Match 37.5%; Score 33; DB 22; Length 9;

Best Local Similarity 75.0%; Pred. No. 7.8e+05;

Matches: 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 5 SGNCGLYY 12  
 II III I  
 Db 1 SGRGGLGY 8

## RESULT 4

AAU24423  
 ID AAU24423 standard; Peptide; 9 AA.  
 AC AAU24423;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human MHC molecule HLA-B35 binding 103P2D6 peptide #8.  
 XX  
 KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;  
 KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;  
 KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;  
 KW single chain monoclonal antibody; cervix; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200162925-A2.  
 XX  
 PD 30-AUG-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US05996.  
 XX  
 PR 24-FEB-2000; 2000US-0184558.  
 PR 13-JUL-2000; 2000US-0218856.  
 XX  
 PA (UROC-) UROGENESYS INC.  
 XX  
 PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;  
 PI Challita-eid PM, Faris M, Jakobovits A;  
 XX  
 DR WPI; 2001-557705/62.  
 XX  
 PT New polynucleotide for treating and diagnosing prostate cancer is the  
 PT 103P2D6 gene which encodes for 103P2D6-related proteins -  
 XX  
 PS Example 15; Page 97; 132pp; English.  
 XX  
 CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and  
 CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal  
 CC adult tissue but is aberrantly expressed in some foetal tissues and many  
 CC cancers including tumours of the prostate, testis, bladder, bone, cervix,  
 CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its  
 CC related protein and also peptide fragments of the protein are therefore  
 CC useful for diagnosing and treating cancer. A vector comprising a  
 CC polynucleotide which encodes a single chain monoclonal antibody, that  
 CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme  
 CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,  
 CC are both useful in the preparation of a composition for treating a  
 CC patient with a cancer that expresses 103P2D6. The sequences can be used  
 CC in diagnostic methods to monitor the level of 103P2D6 gene products in  
 CC serum, blood, urine and tissue and to thereby detect the presence of  
 CC cancerous cells.  
 XX  
 SQ Sequence 9 AA;

Query Match 37.5%; Score 33; DB 22; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLYY 12  
 II III I  
 Db 2 SGRGGLGY 9

## RESULT 5

AAU24441  
 ID AAU24441 standard; Peptide; 9 AA.  
 XX  
 AC AAU24441;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human MHC molecule HLA-B35 binding 103P2D6 peptide #26.  
 XX  
 KW 103P2D6; PCR primer; DNA adaptor; prostate; testis; foetal tissue;  
 KW tumour; cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;  
 KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;  
 KW single chain monoclonal antibody; cervix; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200162925-A2.  
 XX  
 PD 30-AUG-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US05996.  
 XX  
 PR 24-FEB-2000; 2000US-0184558.  
 PR 13-JUL-2000; 2000US-0218856.  
 XX  
 PA (UROC-) UROGENESYS INC.  
 XX  
 PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;  
 PI Challita-eid PM, Faris M, Jakobovits A;  
 XX  
 DR WPI; 2001-557705/62.  
 XX  
 PT New polynucleotide for treating and diagnosing prostate cancer is the  
 PT 103P2D6 gene which encodes for 103P2D6-related proteins -  
 XX  
 PS Example 15; Page 98; 132pp; English.  
 XX  
 CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and  
 CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal  
 CC adult tissue but is aberrantly expressed in some foetal tissues and many  
 CC cancers including tumours of the prostate, testis, bladder, bone, cervix,  
 CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its  
 CC related protein and also peptide fragments of the protein are therefore  
 CC useful for diagnosing and treating cancer. A vector comprising a  
 CC polynucleotide which encodes a single chain monoclonal antibody, that  
 CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme  
 CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,  
 CC are both useful in the preparation of a composition for treating a  
 CC patient with a cancer that expresses 103P2D6. The sequences can be used  
 CC in diagnostic methods to monitor the level of 103P2D6 gene products in  
 CC serum, blood, urine and tissue and to thereby detect the presence of  
 CC cancerous cells.  
 XX  
 SQ Sequence 9 AA;

Query Match 37.5%; Score 33; DB 22; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 SGNCGLYY 12  
 II III I  
 Db 1 SGRGGLGY 8

## RESULT 6

AAU24484  
 ID AAU24484 standard; Peptide; 10 AA.  
 XX  
 AC AAU24484;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE Human MHC molecule HLA-A35 binding 103P2D6 peptide #19.

XX 103P2D6: PCR primer: DNA adaptor: prostate: testis: foetal tissue;  
 KW tumour: cancer; bone; ovary; breast; pancreas; colon; lung; cytostatic;  
 KW gene therapy; antibody therapy; ribozyme; serum; blood; urine; bladder;  
 KW single chain monoclonal antibody; cervix; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200162925-A2.  
 XX  
 PD 30-AUG-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US05996.  
 XX  
 PR 24-FEB-2000; 2000US-0184558.  
 PR 13-JUL-2000; 2000US-0218856.  
 XX  
 PA (UROG-) UROGENESYS INC.  
 XX  
 PI Raitano AB, Afar DEH, Rastegar GS, Mitchell SC, Hubert RS;  
 PI Chailita-eid PM, Faris M, Jakobovits A;  
 XX  
 DR WPI: 2001-557705/62.  
 XX  
 PT New polynucleotide for treating and diagnosing prostate cancer is the  
 PT 103P2D6 gene which encodes for 103P2D6-related proteins.  
 XX  
 PS Example 15; Page 99; 132pp; English.  
 XX  
 CC Sequences AAU23815-AAU24515 represent the 103P2D6-related protein and  
 CC peptide fragments of the polypeptide. 103P2D6 is not expressed in normal  
 CC adult tissue but is aberrantly expressed in some foetal tissues and many  
 CC cancers including tumours of the prostate, testis, bladder, bone, cervix,  
 CC ovary, breast, pancreas, colon and lung. The 103P2D6 polynucleotide, its  
 CC related protein and also peptide fragments of the protein are therefore  
 CC useful for diagnosing and treating cancer. A vector comprising a  
 CC polynucleotide which encodes a single chain monoclonal antibody, that  
 CC immunospecifically binds to an 103P2D6-related protein, and a ribozyme  
 CC capable of cleaving a polynucleotide having the 103P2D6 coding sequence,  
 CC are both useful in the preparation of a composition for treating a  
 CC patient with a cancer that expresses 103P2D6. The sequences can be used  
 CC in diagnostic methods to monitor the level of 103P2D6 gene products in  
 CC serum, blood, urine and tissue and to thereby detect the presence of  
 CC cancerous cells.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 37.5%; Score 33; DB 22; Length 10;  
 Best Local Similarity 75.0%; Pred. NO. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 5 SGNGLYX 12  
 |||||  
 Db 2 SGRGLGY 9  
 RESULT 7  
 AAB26044  
 ID AAB26044 standard; Peptide; 12 AA.  
 XX  
 AC AAB26044;  
 XX  
 XX 05-JAN-2001 (first entry)  
 DT  
 XX Human IgE C-epsilon-2 PtmAb0011 peptide ligand SEQ ID NO:138.  
 DE  
 XX Epitope: minotope; human: immunoglobulin E; IgE; C-epsilon-2 domain;  
 KW allergic disease; immunoprophylaxis; immunotherapy; antiallergic;  
 KW immunosuppressive; vaccine; histamine release inhibitor; immunogen;  
 KW allergy; atopy.  
 XX  
 OS Homo sapiens.  
 XX

PN WO2000050460-A1.  
 XX  
 PD 31-AUG-2000.  
 XX  
 PF 22-FEB-2000; 2000WO-EP01455.  
 XX  
 PR 25-FEB-1999; 99GB-0004405.  
 PR 29-MAR-1999; 99GB-0007151.  
 PR 07-MAY-1999; 99GB-0010537.  
 PR 07-MAY-1999; 99GB-0010538.  
 PR 07-AUG-1999; 99GB-0018594.  
 PR 07-AUG-1999; 99GB-0018603.  
 PR 07-SEP-1999; 99GB-0021046.  
 PR 07-SEP-1999; 99GB-0021047.  
 PR 29-OCT-1999; 99GB-0025619.  
 PR 23-NOV-1999; 99GB-0027698.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PA (PEPT-) PEPTIDE THERAPEUTICS LTD.  
 XX  
 PI Dyson M, Friede M, Greenwood J, Hewitt E, Lamont A, Mason S;  
 PI Randall R, Turnell WG, Van Mechelen MP, Vinals De Bassols YC;  
 XX  
 DR WPI: 2000-572073/53.  
 XX  
 PT Peptides useful for treating, preventing and ameliorating allergic  
 PT diseases, comprising an isolated surface exposed group of a specific  
 PT domain from immunoglobulin E -  
 XX  
 PS Example 7; Page 53; 129pp; English.  
 CC The present invention describes a peptide (I) comprising an isolated  
 CC surface exposed group/epitope (El) of C-epsilon-2 domain (D) of  
 CC immunoglobulin E (IgE) or its minotope, also described are: (1) an  
 CC immunogen (II) for treating allergy comprising (1); (2) a vaccine (III)  
 CC for treating allergies comprising (II); (3) a ligand (IV) capable of  
 CC recognising El; (4) a pharmaceutical composition (PC) comprising (IV);  
 CC (5) a peptide (Ia) capable of being recognised by (IV); (6) an immunogen  
 CC (IIa) comprising (Ia); and (7) producing (III) by producing (II). (I)  
 CC can have anti-allergic and immunosuppressive activities, and can be used  
 CC as a vaccine and histamine release inhibitor. (I), (II) and (III) are  
 CC useful in medicine and in the manufacture of medicaments for treating  
 CC and preventing allergies. (IV) is useful for identifying mimotopes of Pl,  
 CC in medicine and also in manufacturing medicaments for treating  
 CC allergies. (I) is useful in diagnostics and in the affinity purification  
 CC of circulating anti-IgE antibodies from blood. (I), (III) and PC are  
 CC useful for treating a patient susceptible to or suffering from allergies.  
 CC (IV) is also useful in diagnosing atopy. AAB25907 to AAB26099 represent  
 CC peptide sequences which are used in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 37.5%; Score 33; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. NO. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 PSGNCG 9  
 |||||  
 Db 4 PSGDCG 9  
 RESULT 8  
 AAU16769  
 ID AAU16769 standard; Peptide; 12 AA.  
 XX  
 AC AAU16769;  
 XX  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE Peptide EED18/47/48 derived from Cepsilon2 region of human IgE.  
 XX  
 KW Human; linkage technology; conjugated compound; carrier vehicle;

KW epitope; Cepsilon2; Cepsilon3; Cepsilon4; immunoglobulin E;  
 KW IgE mediated disease; antibody response.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200145745-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 21-DEC-2000; 2000WO-CB04935.  
 XX  
 PR 21-DEC-1999; 99GB-0030233.  
 PR 22-FEB-2000; 2000GB-0004096.  
 PR 22-AUG-2000; 2000GB-0020707.  
 PR 22-AUG-2000; 2000GB-0020708.  
 XX  
 PA (ACAM-) ACAMBIS RES LTD.  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 Flinn N, Johnson T;  
 WPI; 2001-521967/57.  
 XX  
 PT A linkage comprising an immunogenic conjugate useful treatment of IgE  
 PT mediated diseases.  
 XX  
 PS Example 4; Page 23; 48pp; English.  
 XX  
 CC The present invention relates to linkage methodology for use in the  
 CC conjugation of compounds (e.g. peptides) to carrier vehicles  
 CC (e.g. macromolecules, polymers, dendrimers, proteins) to produce  
 CC biological and immunological constructs. The invention provides a  
 CC method for linking an epitope (e.g. a peptide) to a carrier (e.g. a  
 CC protein) for use in a pharmaceutical composition or a vaccine. The  
 CC invention describes peptides derived from or mimotopes of the  
 CC Cepsilon2, Cepsilon3 or Cepsilon4 regions of human immunoglobulin E  
 CC (IgE) which are used to produce conjugated compounds. The compounds or  
 CC compositions of the invention are useful in the manufacture of a  
 CC medicament for the treatment of IgE mediated diseases. The invention  
 CC allows for controlled conjugation of a peptide epitope (antigen) to a  
 CC protein so as to form an immunogenic conjugate which may be able to  
 CC raise a protective antibody response in an animal or human patient.  
 CC AAU16632-AAU16913 represent peptides derived from or mimotopes of  
 CC the Cepsilon2/Cepsilon3/Cepsilon4 region of human IgE.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 37.5%; Score 33; DB 22; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 PSGNCG 9  
 Db 4 PSGDCG 9  
 RESULT 9  
 ABJ00390  
 ID ABJ00390 standard; Peptide; 12 AA.  
 XX  
 AC ABJ00390;  
 XX  
 XX 02-SEP-2002 (first entry)  
 DT Human IgE cyclic immunogenic peptide SEQ ID NO: 174.  
 DE Immunogen; human; IgE; immunoglobulin E; allergy; thio-ether linkage;  
 KW vaccine; antiallergic; cyclic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200216409-A2.

XX 28-FEB-2002.  
 XX  
 PF 17-AUG-2001; 2001WO-EP09576.  
 XX  
 PR 22-AUG-2000; 2000GB-0020717.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 PA (PEPT-) PEPTIDE THERAPEUTICS LTD.  
 XX  
 PI Friede M, Mason S, Turnell WG, Vinals Bassols YC;  
 XX  
 XX WPI; 2002-489648/52.  
 DR  
 XX Conjugate for use in vaccine for treatment of allergy, comprises  
 XX disulfide bridge cyclized peptide and immunogenic carrier -  
 PT  
 PS Claim 4; Page 13; 45pp; English.  
 XX  
 CC The present invention relates to conjugates suitable for use in vaccines,  
 CC where the conjugate comprises a disulphide bridge cyclised peptide and an  
 CC immunogenic carrier. The vaccines can be used in the treatment of  
 CC allergies. The present sequence is a cyclic peptide immunogen derived  
 CC from human immunoglobulin E (IgE) suitable for use in the invention.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 37.5%; Score 33; DB 23; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 PSGNCG 9  
 Db 4 PSGDCG 9  
 RESULT 10  
 AAW23509  
 ID AAW23509 standard; peptide; 15 AA.  
 XX  
 AC AAW23509;  
 XX  
 DT 18-SEP-1997 (first entry)  
 XX  
 DE Purified cis-9,10-octadecenoamide internal amino acid fragment.  
 XX  
 KW Coase; cis-9,10-octadecenoamide; oleic acid; catalysis;  
 KW affinity chromatography; electric chromatography;  
 KW gel filtration chromatography; ion exchange chromatography;  
 KW partition chromatography; fatty acid primary amide; sleep-inducing;  
 KW inhibitor; soporific; rat.  
 XX  
 OS Rattus rattus.  
 XX  
 PN WO9641869-A1.  
 XX  
 PD 27-DEC-1996.  
 XX  
 PF 12-JUN-1996; 96WO-US10435.  
 XX  
 PR 12-JUN-1995; 95US-0489535.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 XX  
 PI Cravatt BF, Gilula NB, Lerner RA;  
 XX  
 DR WPI; 1997-065456/06.  
 XX  
 PT Purified cis-9,10-octadecenoamide - useful for hydrolysing  
 PT sleep-inducing fatty acid primary amide(s), and identifying  
 PT inhibitors  
 XX  
 PS Claim 1; Page 78; 101pp; English.

XX A purified form of cis-9,10-octadecenoamidase (Coase) has been  
 CC obtained by a chromatographic methodology selected from affinity,  
 CC electric, gel filtration, ion exchange and partition chromatography.  
 CC The Coase is characterised by enzymic activity for catalysing the  
 CC conversion of cis-9,10-octadecenoamide (CO) to oleic acid and by the  
 CC inclusion of an amino acid sequence fragment from rat liver Coase.  
 CC The present sequence represents a specifically claimed example of such  
 CC an amino acid fragment. This fragment is found at residues 31 to 45 of  
 CC rat liver Coase as shown in AAW10465. The Coase can be used to catalyse  
 CC the hydrolysis of fatty acid primary amides, which have sleep-inducing  
 CC activity. The Coase can also be used to identify inhibitors of the  
 CC Coase activity.

XX SQ Sequence 15 AA;  
 Query Match 37.5%; Score 33; DB 18; Length 15;  
 Best Local Similarity 55.6%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Y 2 RYPSGNCGL 10  
 I:|||||  
 Db 1 RPSAFCGI 9

RESULT 11  
 AAW57788  
 ID AAW57788 standard; Peptide; 15 AA.  
 AC AAW57788;  
 XX AAW57788;  
 DT 12-OCT-1998 (first entry)  
 DE Fatty acid amide hydrolase peptide (e).  
 KW Cis-9,10-octadecenoamidase; fatty acid amide hydrolase; FAAH;  
 KW oleamide hydrolase; sporific; sleep.  
 XX Mammalia.  
 OS  
 XX WO9820119-A1.  
 PD 14-MAY-1998.  
 XX  
 PF 04-NOV-1997; 97WO-US20385.  
 XX  
 PR 04-NOV-1996; 96US-0743168.  
 XX  
 A (SCRI ) SCRIPPS RES INST.  
 PI Cravatt BF, Gilula NB, Lerner RA;  
 XX WPT; 1998-286935/25.  
 DR  
 XX New fatty acid amide hydrolase hydrolysing sporific unsaturated  
 PT amide(s) - useful for, e.g. studying, and potentially developing  
 PT agents for modulating sleep processes  
 XX  
 PS Claim 5; Page 111; 151pp; English.

XX Fatty acid amide hydrolases (FAAHs) of the invention are  
 CC characterised by inclusion of an amino acid sequence selected from  
 CC a group of 28 sequences (see AAW57784-811). These FAAHs can hydrolyse  
 CC cis-9,10-octadecenoamide, anandamide (arachidoyl ethanolamine), and  
 CC myristic, palmitic or stearic amides. Also new are: (1) Inhibiting  
 CC hydrolysis of fatty acid primary amides catalysed by FAAH by  
 CC treatment with an FAAH inhibitor; (2) a method of screening for an  
 CC FAAH inhibitor; (3) the FAAH inhibitor of formula  
 CC CF3CO(CH2)7-CH=CH(cis)-(CH2)7Me1; and (4) nucleic acid encoding FAAH  
 CC or parts of it. FAAH catalyses conversion of fatty acid primary  
 CC amides, particularly those in which the alkyl group has a cis  
 CC unsaturation. These amides induce sleep, so FAAH can be used to  
 CC study processes in which they are involved and also to develop

CC agents for modulating sleep.  
 XX  
 SQ Sequence 15 AA;  
 Query Match 37.5%; Score 33; DB 19; Length 15;  
 Best Local Similarity 55.6%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RYPSGNCGL 10  
 I:|||||  
 Db 1 RPSAFCGI 9

RESULT 12  
 AAG96214  
 ID AAG96214 standard; Peptide; 10 AA.  
 XX AAG96214;  
 AC AAG96214;  
 XX AAG96214;  
 DT 18-SEP-2001 (first entry)  
 DE Human complementary peptide, SEQ ID NO: 2408.  
 DE Human; complementary peptide; ligand; drug discovery; drug design.  
 KW Homo sapiens.  
 OS  
 XX WO200142277-A2.  
 PN 14-JUN-2001.  
 PD 13-DEC-2000; 2000WO-GB04776.  
 PF 13-DEC-1999; 99GB-0029464.  
 PR (PROT-) PROTEOM LTD.  
 XX Roberts GW, Heal JR;  
 PI WPI; 2001-408419/43.  
 DR  
 XX A set of peptide ligands consisting of specific complementary peptides  
 PT to proteins encoded by genes of the human genome, useful in an assay  
 PT for screening and identifying of one or more novel peptides which are  
 PT drug candidates or pro-drugs -  
 XX Example 4; Page 386; 646pp; English.  
 PS  
 CC The invention relates to a set of complementary peptide ligands  
 CC generated from the human genome. The complementary peptides  
 CC interact with their relevant target proteins encoded in the human  
 CC genome. They can be used as reagents in drug discovery and as lead  
 CC ligands to facilitate drug design and development. The present  
 CC sequence is a complementary peptide provided in the specification.

XX SQ Sequence 10 AA;  
 Query Match 35.2%; Score 31; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SGNGC 9  
 I:|||||  
 Db 6 SGNGC 10

RESULT 13  
 AAP40328  
 ID AAP40328 standard; peptide; 15 AA.  
 XX AAP40328;  
 AC AAP40328;  
 XX 16-AUG-2002 (updated)  
 DT



XX Identifying nucleic acid encoding A-lineage conotoxin peptide(s) by  
PT amplification - uses primers corresponding to conserved regions in  
PT the signal sequence and 3'-untranslated regions, useful e.g. in  
PT treatment of small cell lung cancer  
XX  
PS Disclosure; Column 3; 36pp; English.  
PS  
XX  
CC AAW12726-W12769 represent conotoxin peptides. This sequence represents  
CC the A-lineage conotoxin peptide isolated from Conus geographus. These  
CC sequences are identified using the method of the invention. The method  
CC of the invention is for identifying DNA encoding A-lineage conotoxin  
CC peptides by subjecting Conus nucleic acid to amplification with primer  
CC sequences (see AAT59714 and AAT59715). The primers are specific for the  
CC signal sequence and 3'-untranslated (3'UTR) regions of the conotoxin  
CC gene, which are highly homologous between conotoxins, and are therefore  
CC suitable sites for detection. A-lineage conotoxins include alpha-  
CC conotoxins, and kappa-conotoxins. Alpha-conotoxins are powerful  
CC inhibitors of synaptic transmission at the neuromuscular junction, and  
CC are usually nicotinic acetylcholine receptor blockers. Kappa-conotoxins  
CC act on the voltage sensitive sodium and potassium channels. The  
CC conotoxins identified can be used as muscle relaxants, in the diagnosis  
CC of myasthenia gravis, and for the treatment or diagnosis of small cell  
CC lung cancer. For the treatment of small cell lung cancer, the conotoxin  
CC peptides act by binding to the nicotinic receptors, and thereby blocking  
CC the nicotine/cytosine stimulated release of the mitogen  
CC 5-hydroxytryptamine.  
XX  
SQ Sequence 15 AA;

Query Match 35.2%; Score 31; DB 18; Length 15;  
Best Local Similarity 62.5%; Pred. No. 4.1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYYSSG 15  
||:|:|  
Db 7 GGRHYSCG 14

Search completed: December 2, 2002, 07:01:44  
Job time : 92 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein --protein search, using sw model

Run on: December 2, 2002, 07:02:41 ; Search time 10 seconds  
(without alignments)  
23.886 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRVPSNCGLYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 102317 seqs, 15924203 residues

Total number of hits satisfying chosen parameters: 26705

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*

- 1: /cgn2.6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2.6/ptodata/1/pubpaa/PTCT\_NEW\_PUB.pep.\*
- 3: /cgn2.6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2.6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2.6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2.6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2.6/ptodata/1/pubpaa/PTCTUS\_PUBCOMB.pep.\*
- 8: /cgn2.6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2.6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2.6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2.6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2.6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2.6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*
- 14: /cgn2.6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 35    | 39.8        | 10     | 10    | US-09-819-308-25   |
| 2          | 27    | 30.7        | 9      | 10    | US-09-832-723-80   |
| 3          | 27    | 30.7        | 15     | 9     | US-10-086-623-25   |
| 4          | 26.5  | 30.1        | 9      | 10    | US-09-894-018-264  |
| 5          | 26    | 29.5        | 10     | 10    | US-09-765-086-32   |
| 6          | 26    | 29.5        | 10     | 10    | US-09-765-086-55   |
| 7          | 26    | 29.5        | 12     | 9     | US-09-252-150-74   |
| 8          | 26    | 29.5        | 12     | 10    | US-09-879-257A-46  |
| 9          | 26    | 29.5        | 14     | 10    | US-09-758-008-14   |
| 10         | 26    | 29.5        | 14     | 10    | US-09-758-008-15   |
| 11         | 26    | 29.5        | 15     | 8     | US-08-424-550B-551 |
| 12         | 25    | 28.4        | 6      | 10    | US-09-903-452-13   |
| 13         | 25    | 28.4        | 12     | 9     | US-09-880-132-51   |
| 14         | 25    | 28.4        | 12     | 10    | US-09-880-149-51   |
| 15         | 25    | 28.4        | 14     | 10    | US-09-873-637-5    |
| 16         | 25    | 28.4        | 15     | 10    | US-09-867-852-16   |
| 17         | 24    | 27.3        | 9      | 10    | US-09-780-053-447  |
| 18         | 24    | 27.3        | 9      | 10    | US-09-780-053-527  |
| 19         | 24    | 27.3        | 9      | 10    | US-09-780-053-634  |

|    |    |      |    |    |                   |                    |
|----|----|------|----|----|-------------------|--------------------|
| 20 | 24 | 27.3 | 10 | 10 | US-09-758-318-42  | Sequence 42, Appl  |
| 21 | 24 | 27.3 | 10 | 10 | US-09-780-053-471 | Sequence 471, Appl |
| 22 | 24 | 27.3 | 10 | 10 | US-09-780-053-716 | Sequence 716, Appl |
| 23 | 24 | 27.3 | 13 | 10 | US-09-056-160B-76 | Sequence 76, Appl  |
| 24 | 24 | 27.3 | 14 | 10 | US-09-758-008-13  | Sequence 13, Appl  |
| 25 | 24 | 27.3 | 15 | 10 | US-09-767-460-7   | Sequence 7, Appl   |
| 26 | 24 | 27.3 | 15 | 10 | US-09-953-510-47  | Sequence 47, Appl  |
| 27 | 24 | 27.3 | 15 | 10 | US-09-953-510-48  | Sequence 48, Appl  |
| 28 | 24 | 27.3 | 15 | 10 | US-09-953-510-49  | Sequence 49, Appl  |
| 29 | 23 | 26.1 | 7  | 10 | US-09-484-704-14  | Sequence 14, Appl  |
| 30 | 23 | 26.1 | 11 | 10 | US-09-881-276-17  | Sequence 17, Appl  |
| 31 | 23 | 26.1 | 12 | 9  | US-09-826-290-96  | Sequence 96, Appl  |
| 32 | 23 | 26.1 | 12 | 10 | US-09-832-312-75  | Sequence 75, Appl  |
| 33 | 23 | 26.1 | 12 | 10 | US-09-791-378-148 | Sequence 148, Appl |
| 34 | 23 | 26.1 | 14 | 10 | US-09-904-599A-6  | Sequence 6, Appl   |
| 35 | 23 | 26.1 | 14 | 10 | US-09-966-147-37  | Sequence 37, Appl  |
| 36 | 23 | 26.1 | 15 | 10 | US-09-767-460-33  | Sequence 33, Appl  |
| 37 | 22 | 25.0 | 7  | 10 | US-09-823-444-5   | Sequence 5, Appl   |
| 38 | 22 | 25.0 | 8  | 10 | US-09-809-517A-22 | Sequence 22, Appl  |
| 39 | 22 | 25.0 | 8  | 10 | US-09-756-283A-54 | Sequence 54, Appl  |
| 40 | 22 | 25.0 | 9  | 10 | US-09-988-019-23  | Sequence 23, Appl  |
| 41 | 22 | 25.0 | 9  | 10 | US-09-859-214-11  | Sequence 11, Appl  |
| 42 | 22 | 25.0 | 9  | 10 | US-09-791-378-569 | Sequence 569, Appl |
| 43 | 22 | 25.0 | 9  | 10 | US-09-773-308-82  | Sequence 82, Appl  |
| 44 | 22 | 25.0 | 9  | 10 | US-09-773-308-206 | Sequence 206, Appl |
| 45 | 22 | 25.0 | 9  | 10 | US-09-773-308-283 | Sequence 283, Appl |

#### ALIGNMENTS

##### RESULT 1

US-09-819-308-25  
; Sequence 25, Application US/09819308  
; Patent No. US20020019040A1  
; GENERAL INFORMATION:  
; APPLICANT: No. US20020019040A1Aleborn, Mathieu  
; APPLICANT: Dahan-van Oorschot, Astrid  
; APPLICANT: Rohn, Jennifer  
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN  
; FILE REFERENCE: 2906-4820US  
; CURRENT APPLICATION NUMBER: US/09/819,308  
; CURRENT FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 25  
; LENGTH: 10  
; TYPE: PPT  
; ORGANISM: Sequence homology analysis of AAP-5  
US-09-819-308-25

Query Match 39.8%; Score 35; DB 10; Length 10;  
Best Local Similarity 85.7%; Pred. No. 9.6;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 CGLYSS 14

Db 3 CGLYSS 9

##### RESULT 2

US-09-832-723-80  
; Sequence 80, Application US/09832723  
; Patent No. US20020098524A1  
; GENERAL INFORMATION:  
; APPLICANT: Estell, David A.  
; APPLICANT: Chen, Yiyu  
; APPLICANT: Murray, Christopher J.  
; APPLICANT: Tijerina, Pilar  
; TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING  
; FILE REFERENCE: GC617-2  
; CURRENT APPLICATION NUMBER: US/09/832,723  
; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: US 60/197,259  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 117  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 80  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptides screened from a phage display random  
; OTHER INFORMATION: peptide library  
US-09-832-723-80

Query Match 30.7%; Score 27; DB 10; Length 9;  
Best Local Similarity 80.0%; Pred. No. 8.5e+04;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 PSNC 8  
|||:  
DB 5 PSNC 9

RESULT 3  
US-10-086-623-25

; Sequence 25, Application US/10086623  
; Patent No. US20020164710A1

; GENERAL INFORMATION:

; APPLICANT: ERIKSSON, Ulf

; APPLICANT: AASE, Karin

; APPLICANT: Li, Xuri

; APPLICANT: FONTEN, Annica

; APPLICANT: UTELA, Marko

; APPLICANT: ALITALO, Kari

; APPLICANT: OESTMAN, Arne

; APPLICANT: HELDIN, Carl-Henrik

; TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR D, DNA CODING THEREFOR AND USES TH

; FILE REFERENCE: 1064/44833C2

; CURRENT APPLICATION NUMBER: US/10/086,623

; CURRENT FILING DATE: 2000-03-04

; PRIOR APPLICATION NUMBER: US 60/107,852

; PRIOR FILING DATE: 1998-11-10

; PRIOR APPLICATION NUMBER: US 60/113,997

; PRIOR FILING DATE: 1998-12-28

; PRIOR APPLICATION NUMBER: US 60/150,604

; PRIOR FILING DATE: 1999-08-26

; PRIOR APPLICATION NUMBER: US 60/157,108

; PRIOR FILING DATE: 1999-10-04

; PRIOR APPLICATION NUMBER: US 60/157,756

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: US 09/438,046

; PRIOR FILING DATE: 1999-11-10

; PRIOR APPLICATION NUMBER: US 09/691,200

; PRIOR FILING DATE: 2000-10-19

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 15

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: MISC\_FEATURE

; LOCATION: (2)..(2)

; OTHER INFORMATION: can be any amino acid residue

; NAME/KEY: MISC\_FEATURE

; LOCATION: (7)..(7)

; OTHER INFORMATION: can be any amino acid residue

US-10-086-623-25

Query Match 30.7%; Score 27; DB 9; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GNCC 9

DB 11 GNCC 14  
|||:

RESULT 4

US-09-894-018-264

; Sequence 264, Application US/09894018

; Patent No. US20020119127A1

; GENERAL INFORMATION:

; APPLICANT: EPIMUNE, Inc.

; APPLICANT: Sette, Alessandro

; APPLICANT: Chestnut, Robert

; APPLICANT: Livingston, Brian

; APPLICANT: Baker, Denisw

; APPLICANT: Newman, Mark

; APPLICANT: Brown, David

; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING

; FILE REFERENCE: 39963-20033.00

; CURRENT APPLICATION NUMBER: US/09/894,018

; CURRENT FILING DATE: 2001-06-27

; PRIOR APPLICATION NUMBER: PCT/US00/35568

; PRIOR FILING DATE: 2000-12-28

; PRIOR APPLICATION NUMBER: US 60/173,390

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: US 60/284,221

; PRIOR FILING DATE: 2001-04-16

; NUMBER OF SEQ ID NOS: 368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 264

; LENGTH: 9

; TYPE: PRT

; ORGANISM: P. falciparum

US-09-894-018-264

Query Match 30.1%; Score 26.5; DB 10; Length 9;  
Best Local Similarity 66.7%; Pred. No. 8.5e+04;

Matches 6; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 4 PS-GNCC 11

|||:  
DB 1 PS-DGKCNLY 9

RESULT 5

US-09-765-086-32

; Sequence 32, Application US/09765086

; Patent No. US20010046498A1

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki

; APPLICANT: Pasqualini, Renata

; APPLICANT: Wadih, Arap

; APPLICANT: Bredesen, Dale E.

; APPLICANT: Ellerby, H. Michael

; TITLE OF INVENTION: Chimeric Prostate-Homing Peptides With

; FILE REFERENCE: P-LJ 3844

; CURRENT APPLICATION NUMBER: US/09/765,086

; PRIOR FILING DATE: 2001-01-17

; PRIOR APPLICATION NUMBER: US 09/489,582

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 235

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 32

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetic peptide

US-09-765-086-32

Query Match 29.5%; Score 26; DB 10; Length 10;  
Best Local Similarity 66.7%; Pred. No. 2e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 SGNCGL 10  
: | | | |  
Db 5 NGRCGL 10

## RESULT 6

US-09-765-086-55  
; Sequence 55, Application US/09765086  
; Patent No. US20010046498A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; APPLICANT: Wadli, Arap  
; APPLICANT: Bredesen, Dale E.  
; APPLICANT: Ellerby, H. Michael  
; TITLE OF INVENTION: Chimeric prostate-homing peptides with  
; FILE REFERENCE: P-LJ 3844  
; CURRENT APPLICATION NUMBER: US/09765,086  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US 09/489,582  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 235  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 55  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-09-765-086-55

Query Match 29.5%; Score 26; DB 10; Length 10;  
Best Local Similarity 66.7%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 SGNCGL 10  
: | | | |  
Db 5 NGRCGL 10

## RESULT 7

US-09-252-150-74  
; Sequence 74, Application US/09252150A  
; Patent No. US20020155604A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden Ledbetter, Martha  
; APPLICANT: Brady, William A.  
; APPLICANT: Grosmaire, Laura S.  
; APPLICANT: Law, Che-Leung  
; APPLICANT: Dua, Raj  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING  
; FILE REFERENCE: 9113-0019-999  
; CURRENT APPLICATION NUMBER: US/09/252,150A  
; CURRENT FILING DATE: 1999-02-18  
; EARLIER APPLICATION NUMBER: US 60/075,274  
; EARLIER FILING DATE: 1998-02-19  
; EARLIER APPLICATION NUMBER: US 60/108,683  
; EARLIER FILING DATE: 1998-11-16  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 74  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-252-150-74

Query Match 29.5%; Score 26; DB 9; Length 12;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 CGLYYS 13  
: | | | |  
Db 2 CSYYYS 7

## RESULT 8

US-09-879-257A-46  
; Sequence 46, Application US/09879257A  
; Patent No. US20020081690A1  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, SACHIKO  
; APPLICANT: HANADA, TOSHIRO  
; APPLICANT: SHIRO, MINORU  
; APPLICANT: KOBAYAKE, SHINZO  
; TITLE OF INVENTION: HYBRID ENZYMES AND USE THEREOF  
; FILE REFERENCE: 55986(70281)  
; CURRENT APPLICATION NUMBER: US/09/879,257A  
; CURRENT FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Hepatitis B Virus  
US-09-879-257A-46

Query Match 29.5%; Score 26; DB 10; Length 12;  
Best Local Similarity 57.1%; Pred. No. 2.4e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 9 GLYSSG 15  
: | | | : |  
Db 6 GLYFPAG 12

## RESULT 9

US-09-758-008-14  
; Sequence 14, Application US/09758008  
; Patent No. US20020127721A1  
; GENERAL INFORMATION:  
; APPLICANT: Law, Lane K.  
; TITLE OF INVENTION: Adenovirus serotype 30 (Ad30)  
; FILE REFERENCE: 875,044US1  
; CURRENT APPLICATION NUMBER: US/09/758,008  
; CURRENT FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Adenovirus  
US-09-758-008-14

Query Match 29.5%; Score 26; DB 10; Length 14;  
Best Local Similarity 71.4%; Pred. No. 2.8e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PSNCGL 10  
: | | | |  
Db 7 PSPNCL 13

## RESULT 10

US-09-758-008-15  
; Sequence 15, Application US/09758008  
; Patent No. US20020127721A1  
; GENERAL INFORMATION:  
; APPLICANT: Law, Lane K.  
; APPLICANT: Davidson, Beverly L.  
; TITLE OF INVENTION: Adenovirus serotype 30 (Ad30)

FILE REFERENCE: 875.044US1  
CURRENT APPLICATION NUMBER: US/09/758,008  
CURRENT FILING DATE: 2001-01-09  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Adenovirus  
US-09-758-008-15

Query Match 29.5%; Score 26; DB 10; Length 14;  
Best Local Similarity 50.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSGNGLY 11  
|||  
Db 7 PSPNCRIH 14

RESULT 11  
US-08-424-550B-551  
Sequence 551, Application US/08424550B  
Patent No. US20020119447A1  
GENERAL INFORMATION:  
APPLICANT: JOHN N. SIMONS  
APPLICANT: TAMI J. PILOT-MATIAS  
APPLICANT: GEORGE J. DAWSON  
APPLICANT: GEORGE G. SCHLAUDER  
APPLICANT: SURESH M. DESAI  
APPLICANT: THOMAS P. LEARY  
APPLICANT: ANTHONY SCOTT MUERHOFF  
APPLICANT: JAMES C. ERKER  
APPLICANT: SHERI L. BUIJK  
APPLICANT: ISA K. MUSHAWAR  
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
NUMBER OF SEQUENCES: 716  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
STREET: 100 ABBOTT PARK ROAD  
CITY: ABBOTT PARK  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/424,550B  
FILING DATE:  
CLASSIFICATION: 435435  
ATTORNEY/AGENT INFORMATION:  
NAME: FOREMSKI, PRISCILLA E.  
REGISTRATION NUMBER: 33,207  
REFERENCE/DOCKET NUMBER: 5527.PC.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-937-6365  
TELEFAX: 708-938-2623  
INFORMATION FOR SEQ ID NO: 551:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-424-550B-551

Query Match 29.5%; Score 26; DB 8; Length 15;  
Best Local Similarity 66.7%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PSGNCG 9  
|||  
Db 4 PSARCG 9

RESULT 12  
US-09-903-452-13  
Sequence 13, Application US/09903452  
Patent No. US20020042137A1  
GENERAL INFORMATION:  
APPLICANT: Richards, Cynthia  
APPLICANT: Weiner, Michael  
TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES  
FILE REFERENCE: PU3481US2  
CURRENT APPLICATION NUMBER: US/09/903,452  
CURRENT FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: 60/084,936  
PRIOR FILING DATE: 1998-05-11  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 6  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: thrombin  
OTHER INFORMATION: cleavage site  
US-09-903-452-13

Query Match 28.4%; Score 25; DB 10; Length 6;  
Best Local Similarity 80.0%; Pred. No. 8.5e+04;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 YPSGN 7  
|||  
Db 2 YPRGN 6

RESULT 13  
US-09-880-132-51  
Sequence 51, Application US/09880132  
Patent No. US20020173049A1  
GENERAL INFORMATION:  
APPLICANT: Kenten, John  
APPLICANT: Roberts, Steven  
TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS  
FILE REFERENCE: 2757-6  
CURRENT APPLICATION NUMBER: US/09/880,132  
CURRENT FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: 09/406,781  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: 60/119,851  
PRIOR FILING DATE: 1999-02-12  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 51  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: ubiquitination  
OTHER INFORMATION: recognition element  
US-09-880-132-51

Query Match 28.4%; Score 25; DB 9; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SGNC 8  
||||  
Db 9 SGNC 12

## RESULT 14

US-09-880-149-51  
; Sequence 51, Application US/09880149  
; Patent No. US20020146843A1  
; GENERAL INFORMATION:  
; APPLICANT: Kenten, John  
; TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS  
; FILE REFERENCE: 2757-5  
; CURRENT APPLICATION NUMBER: US/09/880,149  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 09/406,781  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: 60/119,851  
; PRIOR FILING DATE: 1999-02-12  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 51  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: ubiquitination  
; OTHER INFORMATION: recognition element  
; US-09-880-149-51

Query Match 28.4%; Score 25; DB 10; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 SGNC 8  
DB 9 SGNC 12

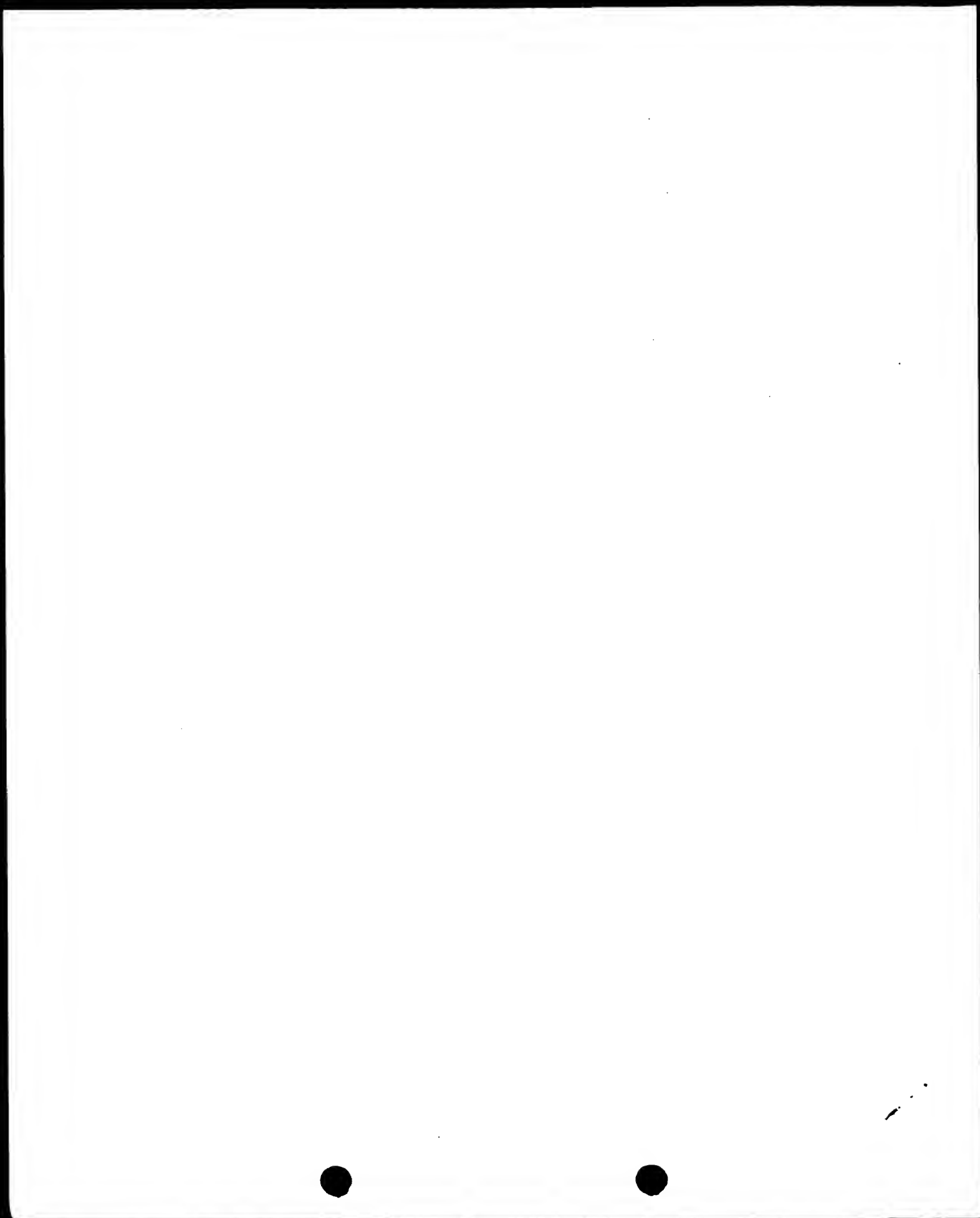
## RESULT 15

US-09-873-637-5  
; Sequence 5, Application US/09873637  
; Patent No. US20020061543A1  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Jeffrey  
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN  
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE  
; FILE REFERENCE: 960296.95131  
; CURRENT APPLICATION NUMBER: US/09/873,637  
; CURRENT FILING DATE: 2001-06-04  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-873-637-5

Query Match 28.4%; Score 25; DB 10; Length 14;  
Best Local Similarity 55.6%; Pred. No. 3.9e+02;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 DRYPSGNCG 9  
DB 6 DRMPGGRGG 14

Search completed: December 2, 2002, 07:10:54  
Job time : 11 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 2, 2002, 06:59:51 ; Search time 14 seconds  
(without alignments)  
31.525 Million cell updates/sec

Title: US-09-902-563-18

Perfect score: 88

Sequence: 1 DRYPSGNCGLYSSG 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

al number of hits satisfying chosen parameters: 111443

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA.\*  
1: /cgn2.6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2.6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2.6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2.6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2.6/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 88    | 100.0       | 15     | 4  | US-09-442-143A-18 |
| 2          | 33    | 37.5        | 15     | 4  | US-08-743-168B-9  |
| 3          | 33    | 37.5        | 15     | 5  | PCT-US96-10435-9  |
| 4          | 31    | 35.2        | 15     | 1  | US-08-137-800-5   |
| 5          | 31    | 35.2        | 15     | 1  | US-08-477-383-5   |
| 6          | 31    | 35.2        | 15     | 1  | US-08-487-174-5   |
| 7          | 31    | 35.2        | 15     | 1  | US-08-480-750-5   |
| 8          | 31    | 35.2        | 15     | 5  | PCT-US96-07962-2  |
| 9          | 30    | 34.1        | 13     | 1  | US-07-923-724-57  |
| 10         | 30    | 34.1        | 13     | 2  | US-08-609-426A-57 |
| 11         | 30    | 34.1        | 13     | 2  | US-08-374-652C-44 |
| 12         | 30    | 34.1        | 15     | 2  | US-08-553-257A-42 |
| 13         | 29    | 33.0        | 15     | 2  | US-08-432-871C-91 |
| 14         | 29    | 33.0        | 15     | 4  | US-09-270-956-91  |
| 15         | 28.5  | 32.4        | 10     | 2  | US-08-318-856A-72 |
| 16         | 28    | 31.8        | 7      | 4  | US-08-492-411A-20 |
| 17         | 28    | 31.8        | 11     | 1  | US-08-466-468-6   |
| 18         | 28    | 31.8        | 11     | 4  | US-08-468-408-6   |
| 19         | 28    | 31.8        | 11     | 4  | US-08-937-228-6   |
| 20         | 28    | 31.8        | 11     | 4  | US-08-492-411A-33 |
| 21         | 28    | 31.8        | 11     | 4  | US-08-421-583-6   |
| 22         | 28    | 31.8        | 11     | 4  | US-09-639-242A-6  |
| 23         | 28    | 31.8        | 12     | 1  | US-08-190-788A-21 |
| 24         | 28    | 31.8        | 12     | 1  | US-08-383-474B-26 |
| 25         | 28    | 31.8        | 12     | 1  | US-08-465-391A-21 |
| 26         | 28    | 31.8        | 12     | 2  | US-08-464-538B-21 |
| 27         | 28    | 31.8        | 12     | 2  | US-08-463-076E-65 |

```

28      28      31.8      12      4      US-08-737-841-13      Sequence 13, Appl
29      28      31.8      13      1      US-08-089-994A-20      Sequence 20, Appl
30      28      31.8      13      5      PCT-US94-07605-20      Sequence 20, Appl
31      28      31.8      15      2      US-08-432-871C-100      Sequence 100, Appl
32      28      31.8      15      4      US-09-347-504-55      Sequence 55, Appl
33      28      31.8      15      4      US-09-270-956-100      Sequence 100, Appl
34      27      30.7      13      3      US-08-881-037-88      Sequence 88, Appl
35      27      30.7      14      1      US-08-466-468-7      Sequence 7, Appl
36      27      30.7      14      4      US-08-468-408-7      Sequence 7, Appl
37      27      30.7      14      4      US-08-937-228-7      Sequence 7, Appl
38      27      30.7      14      4      US-08-421-583-7      Sequence 7, Appl
39      27      30.7      14      4      US-09-639-242A-7      Sequence 7, Appl
40      27      30.7      15      2      US-08-432-871C-80      Sequence 80, Appl
41      27      30.7      15      2      US-08-432-871C-83      Sequence 83, Appl
42      27      30.7      15      3      US-08-469-141A-14      Sequence 14, Appl
43      27      30.7      15      4      US-09-347-504-68      Sequence 68, Appl
44      27      30.7      15      4      US-09-347-504-74      Sequence 74, Appl
45      27      30.7      15      4      US-09-270-956-80      Sequence 80, Appl

```

## ALIGNMENTS

```

RESULT 1
US-09-442-143A-18
; Sequence 18, Application US/09442143A
; Patent No. 6403089
; GENERAL INFORMATION:
; APPLICANT: Levy, Gary
; TITLE OF INVENTION: Methods of Modulating Immune Coagulation
; FILE REFERENCE: 9579-14
; CURRENT APPLICATION NUMBER: US/09/442.143A
; CURRENT FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: US 60/046.537
; PRIOR FILING DATE: 1997-05-17
; PRIOR APPLICATION NUMBER: US 60/061.684
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-442-143A-18

```

```

Query Match      100.0%; Score 88; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 DRYPSGNCGLYSSG 15
Db 1 DRYPSGNCGLYSSG 15

```

```

RESULT 2
US-08-743-168B-9
; Sequence 9, Application US/08743168B
; Patent No. 6271015
; GENERAL INFORMATION:
; APPLICANT: Glilula, No. 6271015ton B
; APPLICANT: Cravatt, Benjamin F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: FATTY-ACID AMIDE HYDROLASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute
; STREET: 10550 No. 6271015th Torrey Pines Road
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/743,168B  
FILING DATE: 04-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/489,535  
FILING DATE: 12-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 485.2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-743-168B-9

Query Match 37.5%; Score 33; DB 4; Length 15;  
Best Local Similarity 55.6%; Pred. No. 55;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RYPSGNCGL 10  
|:|:| |:  
Db 1 RPPSAFCGI 9

RESULT 3  
PCT-US96-10435-9  
Sequence 9, Application PC/TUS9610435  
GENERAL INFORMATION:  
APPLICANT: The Scripps Research Institute  
TITLE OF INVENTION: CIS-9, 10-OCTADECENOAMIDASE  
NUMBER OF SEQUENCES: 32  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/10435  
FILING DATE: 12-JUN-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/489,535  
FILING DATE: 12-JUN-1995  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
PCT-US96-10435-9

Query Match 37.5%; Score 33; DB 5; Length 15;  
Best Local Similarity 55.6%; Pred. No. 55;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RYPSGNCGL 10  
|:|:| |:  
Db 1 RPPSAFCGI 9

RESULT 4  
US-08-137-800-5  
Sequence 5, Application US/08137800  
Patent No. 5514774  
GENERAL INFORMATION:  
APPLICANT: Olivera, Baldomero M.  
APPLICANT: Cruz, Lourdes J.  
APPLICANT: Hillyard, David R.  
APPLICANT: McIntosh, J. Michael  
APPLICANT: Santos, Aneurfino D.  
TITLE OF INVENTION: Conotoxin Peptides  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
STREET: 1201 New York Avenue N.W., Suite 1000  
CITY: Washington  
STATE: DC  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,800  
FILING DATE: 19-OCT-1993  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Innen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24260-104763  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Conus geographus  
US-08-137-800-5

Query Match 35.2%; Score 31; DB 1; Length 15;  
Best Local Similarity 62.5%; Pred. No. 11e-02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 CGLYSSG 15  
|:|:| |:  
Db 7 CGRHYSCG 14

RESULT 5  
US-08-477-383-5  
Sequence 5, Application US/08477383  
Patent No. 5589340  
GENERAL INFORMATION:  
APPLICANT: Olivera, Baldomero M.  
APPLICANT: Cruz, Lourdes J.  
APPLICANT: Hillyard, David R.  
APPLICANT: Macintosh, J. Michael  
APPLICANT: Santos, Aneurfino S.  
TITLE OF INVENTION: Conotoxin Peptides  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
STREET: 1201 New York Avenue, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.

ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,383  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/137,800  
FILING DATE: 19-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/084,848  
FILING DATE: 29-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24260-107673  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Conus geographus  
US-08-477-383-5

Query Match 35.2%; Score 31; DB 1; Length 15;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYSSG 15  
||:||||  
Db 7 CGRHYSG 14

RESULT 6  
US-08-487-174-5  
Sequence 5, Application US/08487174  
Patent No. 5595972  
GENERAL INFORMATION:  
APPLICANT: Olivera, Baldomero M.  
APPLICANT: Cruz, Lourdes J.  
APPLICANT: Hillyard, David R.  
APPLICANT: Macintosh, J. Michael  
APPLICANT: Santos, Ameurfino S.  
TITLE OF INVENTION: Conotoxin Peptides  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
STREET: 1201 New York Avenue, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,174  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/137,800  
FILING DATE: 19-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/084,848  
FILING DATE: 29-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24260-107673  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Conus geographus  
US-08-487-174-5

Query Match 35.2%; Score 31; DB 1; Length 15;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYSSG 15  
||:||||  
Db 7 CGRHYSG 14

RESULT 7  
US-08-480-750-5  
Sequence 5, Application US/08480750  
Patent No. 5633347  
GENERAL INFORMATION:  
APPLICANT: Olivera, Baldomero M.  
APPLICANT: Cruz, Lourdes J.  
APPLICANT: Hillyard, David R.  
APPLICANT: Macintosh, J. Michael  
APPLICANT: Santos, Ameurfino S.  
TITLE OF INVENTION: Conotoxin Peptides  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
STREET: 1201 New York Avenue, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,750  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/137,800  
FILING DATE: 19-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/084,848  
FILING DATE: 29-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24260-107673  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810

TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Conus geographus  
US-08-480-750-5

Query Match 35.2%; Score 31; DB 1; Length 15;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYSSG 15  
||:||||  
b 7 CGRHSCG 14

## RESULT 8

PCT-US96-07962-2  
Sequence 2, Application PC/TUS9607962  
GENERAL INFORMATION:  
APPLICANT: University of Utah Research Foundation  
TITLE OF INVENTION: Use of Conotoxin Peptides U002 and M11  
TITLE OF INVENTION: For Treating or Detecting Small-Cell Lung Carcinoma  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
STREET: 1201 New York Avenue, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-WINDOWS  
SOFTWARE: Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/07962  
FILING DATE: 04-JUN-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/487,174  
FILING DATE: 07-JUN-1995

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Conus geographus  
PCT-US96-07962-2

Query Match 35.2%; Score 31; DB 5; Length 15;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CGLYSSG 15  
||:||||  
b 7 CGRHSCG 14

## RESULT 9

US-07-923-724-57  
Sequence 57, Application US/07923724  
Patent No. 5780292  
GENERAL INFORMATION:  
APPLICANT: Nevalainen, Helena K.M.  
PALOHEIMO, Marja T.  
MIETTINEN-OINONEN, Arja S.K.  
APPLICANT: Torkkeli, Tuula K.  
APPLICANT: Cantrell, Michael  
APPLICANT: Piddington, Christopher S.  
APPLICANT: Rambosek, John A.  
APPLICANT: Turunen, Marja K.  
APPLICANT: Fagerstr m, Richard B.  
TITLE OF INVENTION: Production of Phytase Degrading Enzymes  
TITLE OF INVENTION: in Trichoderma  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/923,724  
FILING DATE: 31-JUL-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/496,155  
FILING DATE: 19-MAR-1990

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/044,077  
FILING DATE: 29-APR-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 8610600  
FILING DATE: 30-APR-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimbala, Michele A.  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1050.0240004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
US-07-923-724-57

Query Match 34.1%; Score 30; DB 1; Length 13;  
Best Local Similarity 55.6%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPGNCG 9  
:||||:  
b 4 ERYPSAG 12

## RESULT 10

US-08-609-426A-57  
Sequence 57, Application US/08609426A  
Patent No. 5830733  
GENERAL INFORMATION:  
APPLICANT: Nevalainen, Helena K.M.  
PALOHEIMO, Marja T.  
MIETTINEN-OINONEN, Arja S.K.

APPLICANT: Torckelli, Tuula K.  
APPLICANT: Cantrell, Michael  
APPLICANT: Piddington, Christopher S.  
APPLICANT: Rambosek, John A.  
APPLICANT: Turunen, Marja K.  
APPLICANT: Fagerstr m, Richard B.  
APPLICANT: Houston, Christine S.  
TITLE OF INVENTION: Production of Phytase Degrading Enzymes  
TITLE OF INVENTION: in Trichoderma  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/609,426A  
FILING DATE: 01-MAR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/923,724  
FILING DATE: 31-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/496,155  
FILING DATE: 19-MAR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/044,077  
FILING DATE: 29-APR-1987  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: UK 8610600  
FILING DATE: 30-APR-1986  
NAME: Reed, Grant E.  
REGISTRATION NUMBER: P-41,264  
REFERENCE/DOCKET NUMBER: 1050.0080001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 57:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
US-08-609-426A-57

Query Match 34.1%; Score 30; DB 2; Length 13;  
Best Local Similarity 55.6%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPSGNCG 9  
:||||: |  
Db 4 ERYPSAG 12

RESULT 11  
US-08-374-652C-44  
; Sequence 44, Application US/08374652C  
; Patent No. 5834286  
; GENERAL INFORMATION:  
; APPLICANT: NEVALAINEN, HELENA K.M.  
; APPLICANT: PALOHEIMO, MARJA T.  
; APPLICANT: FAGERSTROM, RICHARD B.  
; APPLICANT: MIETTINEN-OINONEN, ARJA S.  
; APPLICANT: TURUNEN, MARJA K.  
; APPLICANT: RAMBOSEK, JOHN A.  
; APPLICANT: PIDDINGTON, CHRISTOPHER S.

APPLICANT: HOUSTON, CHRISTINE S.  
APPLICANT: CANTRELL, MICHAEL A.  
TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,  
TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVENUE, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/374,652C  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/07058  
FILING DATE: 27-JUL-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/925,401  
FILING DATE: 31-JUL-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: REED, GRANT E.  
REGISTRATION NUMBER: 41,264  
REFERENCE/DOCKET NUMBER: 1050.071001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-374-652C-44

Query Match 34.1%; Score 30; DB 2; Length 13;  
Best Local Similarity 55.6%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 DRYPSGNCG 9  
:||||: |  
Db 4 ERYPSAG 12

RESULT 12  
US-08-553-257A-42  
; Sequence 42, Application US/08553257A  
; Patent No. 5994083  
; GENERAL INFORMATION:  
; APPLICANT: ISTITUTO DI RICERCHE DI BIOLOGIA  
; APPLICANT: MOLECOLARE P. ANGELETTI S.P.A.  
; APPLICANT: FELICI, Franco  
; APPLICANT: LUZZAGO, Alessandra  
; APPLICANT: NICOSIA, Alfredo  
; APPLICANT: MONACI, Paolo  
; APPLICANT: CORTESE, Riccardo  
TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF IMMUNOGENS  
TITLE OF INVENTION: OR DIAGNOSTIC REAGENTS, AND IMMUNOGENS OR  
TITLE OF INVENTION: DIAGNOSTIC REAGENTS THEREBY OBTAINABLE  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Browdy and Nelmark

STREET: 419 Seventh Street N.W. Ste. 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/553,257A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IT94/00054  
FILING DATE: 05-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: RM93A000301  
FILING DATE: 11-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Browdy, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: FELICI-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-5197  
TELEFAX: (202) 737-3528  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: recombinant protein  
HYPOTHETICAL: yes  
FRAGMENT TYPE: internal  
IMMEDIATE SOURCE:  
LIBRARY: of recombinant peptides on phage  
CLONE: phagic  
FEATURE:  
NAME/KEY: polypeptide  
IDENTIFICATION METHOD: selection with specific antibodies  
US-08-553-257A-42

Query Match 34.1%; Score 30; DB 2; Length 15;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

4 PSGNCG 9  
111 11  
9 PSGACG 14

RESULT 13  
US-08-432-871C-91  
Sequence 91, Application US/08432871C  
Patent No. 5877010  
GENERAL INFORMATION:  
APPLICANT: Loeb, Lawrence A.  
APPLICANT: Black, Margaret E.  
TITLE OF INVENTION: THYMIDINE KINASE MUTANTS  
NUMBER OF SEQUENCES: 104  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/432,871C  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 240052.409C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-432-871C-91

Query Match 33.0%; Score 29; DB 2; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 DRYPSGNCGLY 11  
111 11 1  
DB 4 DRHPIGQTSCY 14

RESULT 14  
US-09-270-956-91  
Sequence 91, Application US/09270956  
Patent No. 6451571  
GENERAL INFORMATION:  
APPLICANT: Loeb, Lawrence A.  
APPLICANT: Black, Margaret E.  
TITLE OF INVENTION: THYMIDINE KINASE MUTANTS  
NUMBER OF SEQUENCES: 104  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/270,956  
FILING DATE: 17-MAR-1999  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 240052.409C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-270-956-91

Query Match 33.0%; Score 29; DB 4; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 DRYPSCNGCLY 11  
Db 4 DRHPIGQTSCY 14

RESULT 15  
US-08-318-856A-72  
; Sequence 72, Application US/08318856A  
; Patent No. 5972351  
; GENERAL INFORMATION:  
; APPLICANT: Adrian V.S. Hill, et al.  
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-  
; TITLE OF INVENTION: RESTRICTED CTL EPTOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE  
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
; STREET: 2033 K Street, N.W., Suite 800  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 5.1+  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,856A  
; FILING DATE: October 3, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 92 08 068.8  
; FILING DATE: April 3, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 92 17 704.7  
; FILING DATE: August 20, 1992  
; APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/GB93/00711  
; FILING DATE: April 5, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee Cheng  
; REGISTRATION NUMBER: 40,949  
; REFERENCE/DOCKET NUMBER: 263-PP1R1577US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 721-8200  
; TELEFAX: (202) 721-8250  
; INFORMATION FOR SEQ ID NO: 72:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acid residues  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-318-856A-72

Query Match 32.4%; Score 28.5; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 3 YPS-GNCGLY 11  
Db 1 HPSDGKCNLY 10

Search completed: December 2, 2002, 07:02:57  
Job time : 15 secs

